

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2002, 23:24:39 ; Search time 970.899 Seconds
(without alignments)
15087.634 Million cell updates/sec

Title: US-09-881-556-3
Perfect score: 700
Sequence: 1 gaattggcagcagattttt.....accgcattctacagaaaac 700

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	DB ID	Description
------------	-------------	--------	-------	-------------

RESULT 1
AF178459
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
-REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES

AF178459 81.6 1310 3 AF178459 3 AF178459 1310 bp mRNA linear INV 01-SEP-2000
Cryptosporidium parvum unknown mRNA sequence.

ALIGNMENTS

AF178459 571.4 52.4 7.5 349980 6 AX344574
AF178459 52 7.4 116696 3 PFMAL3P3
AF178459 51.6 7.4 211376 10 AC073883
AF178459 51 7.3 78578 2 PFMAL13P9
AF178459 50.8 7.3 6928 6 AX346425
AF178459 50.6 7.3 15928 3 AB070263
AF178459 50.4 7.2 168698 9 AC068138
AF178459 50.4 7.2 263 11 CNS068HA
AF178459 50.4 7.2 108908 3 PFMAL3P8
AF178459 50.2 7.2 349980 6 AX344552
AF178459 50.2 7.2 349980 6 AX344563
AF178459 50.2 7.2 349980 6 AX344564
AF178459 49.6 7.1 11670 6 AX281283
AF178459 49.4 7.1 5814 6 AX344573
AF178459 49.2 7.0 215532 2 AC098478
AF178459 48.8 7.0 17280 6 AX251529
AF178459 48.8 7.0 115224 9 AL359539
AF178459 48.4 6.9 7040 6 AX251193
AF178459 48.4 6.9 7040 6 AX347400
AF178459 48.4 6.9 7040 6 AX349121
AF178459 48.4 6.9 11046 6 AX344532
AF178459 48.2 6.9 1141 6 AX083744
AF178459 48.2 6.9 9817 6 AX346270
AF178459 48.2 6.9 17389 6 AX346316
AF178459 48.2 6.9 17839 3 AE001384
AF178459 48 6.9 10528 6 AX345237
AF178459 47.8 6.8 349980 6 AX344561
AF178459 47.6 6.8 10039 6 AX346945
AF178459 47.4 6.8 4785 8 POPHMT3B
AF178459 47.2 6.7 5501 6 AX251943
AF178459 47.2 6.7 5501 6 AX281173
AF178459 47.2 6.7 5501 6 AX356430
AF178459 47.2 6.7 9110 6 AX347330
AF178459 47.2 6.7 14867 3 AE001398
AF178459 47.2 6.7 110000 2 PFMAL4P1_2
AF178459 47.2 6.7 175522 2 AC023970
AF178459 47 6.7 6123 6 AX345722
AF178459 47 6.7 8392 6 AX346392
AF178459 47 6.7 17893 6 AX346267
AF178459 47 6.7 40862 6 AX346974
AF178459 47 6.7 73828 2 AC102154
AF178459 46.8 6.7 205429 2 AC005506
AF178459 46.8 6.7 253305 3 PFMAL3P7

Cryptosporidium parvum.
Cryptosporidium parvum
Eukaryota; Alveolata; Apicomplexa; Coccidia; Elmeriida;
Cryptosporidiidae; Cryptosporidium.
1 (bases 1 to 1310)
Simonsen,J.N., Kramer,M., Lowden,C. and Wilkins,J.
Neutralising antigen 2 recognized by human immune sera
Unpublished
2 (bases 1 to 1310)
Simonsen,J.N., Kramer,M., Lowden,C. and Wilkins,J.
Direct Submission
Submitted (16-AUG-1999) Medical Microbiology, University of
Manitoba, 730 William Ave, Winnipeg, MB R3E 0W3, Canada
NCBI staff are still waiting for submitters to provide appropriate
feature information.
Location/Qualifiers

[illegible]

```

FYLVNSYDDVISFEKNPTEYINVDDSPFKLEQINNFDYNNNNNNNNPIENEGFCLV
SUNDGQFVIGKDSCKIYDKNYIFQNKERMOTFINNDYLNQNFIDILNOLKN
KNDLNTLYLYLCHALYELGKHPYIPGLTIQLSAIKFLALFYKKNKLFSEI
PLRKYEQNYQFLDQCDIPHIQIKLKOKFKNNSTNKNOTINTPVNNWTKRDIKYHK
YINLYDKILYSOTITNKQKKIKYT"
23110..24606
/gene="PFC0340w, MAL3P3.4"
CDS
23110..24606
/gene="PFC0340w, MAL3P3.4"
/notes="PFC0340w, (MAL3P3.4), putative DNA polymerase delta
small subunit, len: 499 aa; Similarity to DNA polymerase
delta, small subunits: A.thaliana DNA polymerase delta,
small subunit (F.R:048520) BLAST Score: 345, sum P(3) =
1.3e-58; 35% in 208 aa overlap, predicted using hexexon"
/codon_start=1
/product="putative DNA polymerase delta small subunit"
/protein_id="CAB11105.1"
/db_xref="GI:3649756"
/db_xref="SPTREMBL:O77321"
/translation="MDEVVTNQEVEEIIKHNNKNGEHNKKKKHKKHFEYVNLSE
RFIIPCPTYQFQSHIYSLRIKIMKTLMTIEDIVKKEGNEKKKKNNNDNNND
NINNDNNNDNNNDNNNDNNNNYNNYHILQYLKEIKVNEKQVCIGTIFK
KMLRPSILNVELSILNNTINVTDEIDLEFDEARLKLEGINSDYVIGTLTV
IVKGIGMSGSLYVEEIIYVLPKLDIPNCISNDKFKILFVSGLYINELKSINNISL
LNFILRHGTILYSENLRILIGNSLSDNDEKMDNTIDVFLSLCSIHIDLMP
GDKPDSNLPQQPPNFFKRNFSQCVTPNPFSDININCCMSGSPVHNIS
YSKNNKMALKLIQSRLSPTDGLGCTPFTKNDPFLDNDNTYHIFINGNSKIL
EYQVMQDHKQLPLVCLVPNFHISPKAFILNKLKLEYKILTFQIDEDK"
25880..31015
/gene="PFC0345w, MAL3P3.5"
CDS
25880..31015
/gene="PFC0345w, MAL3P3.5"
/notes="PFC0345w, (MAL3P3.5), Hypothetical protein, len:
1712 aa"
/codon_start=1
/product="hypothetical protein, PFC0345w"
/protein_id="CAB11106.1"
```

Query Match 7.4%; Score 52; DB 3; Length 116696;

Best Local Similarity 44.6%; Pred. No. 1.6;

Matches 247; Conservative 0; Mismatches 305; Indels 2; Gaps 1;

```

QY 111 AAAACATTCTACTAAATTCGAATTAATTTGAGCTGGCTTGATTAATGGTAGTGG 170
Db 103039 AATAATAATATTAAAGAGATGAGCATAAAGAGAAATCAATGAAGAATGATAAGATG 103098

QY 171 CTCGGAGGTGATATTTTATCCTTGATAGGAACACGCTCTTGAAGCTGTAAGTTATTCA 230
Db 103099 AGTGAATGTGATGTATTACAAATAAAGAGATATGATCATATAATAATAATAT 103158

QY 231 ATCGCTTGTTTCTATACAAAAACATGTTTTGAAAAAGATGAAGACATTTGCTTAAA 290
Db 103159 AACAAATATGATGATGATAATAATAATAACAACATGATGATGATAATAATAAT 103218

QY 291 CCCTTTATCGCGCTGAGATAAATGACTTTTGGTCTCGAAAAGACTTACGCAACT 350
Db 103219 AACACAATATGATGATCATATAATAACAACAATAATGATGATGATAATAATAAC 103278

QY 351 CTCCAATCTCTTAATCTGATATTATCTTTTCCCTACATGGAATAAATGCTATTCTTGA 410
Db 103279 AACACAATAATGATGATGATAATAATAAGATCCCATATTAGATGGAGAACTGTCT 103338

QY 411 TATATT--CCAAATTACACAATAATAAATCAACAAAATTTCTCTTGAAAGTTCGGAGAAA 468
Db 103339 TATATTGATGAATAAATCAAAAAAGAAATAAATATATTTCTTAAATATATAGATAAAAA 103398

QY 469 ACTCAATCTCCCAATTTGATATTGGAGTATTATCGATGGATTTTCTTCTTTTAATA 528
Db 103399 ATTGAATGTTGAACGAATGTTATTATGAAAAATAACGTTTGTGATGGTATTATTAT 103458

QY 529 AACATCGCGCTATAAAGAAAAATGCCACTTTGAATAATCAAGAAATCAAAATATTCAAAAT 588
Db 103459 TTAATTAGACAGAAAATGGTTTGAATATTGTCAGAGAGTTAATAACATATATATAA 103518
```

```

QY 589 GAAATAATGAAGCCACTAACTCAACAAATCCAGATCAGTGGGAATAACTTAACCTCAG 648
Db 103519 AACAAAATGAAGCAAGAAAGTTTATATAAAAAAACAATAAATAATATCAAG 103578

QY 649 AACCAAAACACAAA 662
Db 103579 AACAGAACAGAA 103592

RESULT 4
AC073883/c 211376 bp DNA linear ROD 12-FEB-2002
LOCUS Mus musculus chromosome 3 clone rp23-71el3, complete sequence.
DEFINITION AC073883
ACCESSION AC073883
VERSION HTG.
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 211376)
AUTHORS Hu, X., Swank, R. and Roe, B. A.
TITLE Mus musculus Chromosome 3 BAC Clone rp23-71el3
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 211376)
AUTHORS Hu, X., Swank, R. and Roe, B. A.
TITLE Direct Submission
JOURNAL Submitted (01-JUL-2000) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Farrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 3 (bases 1 to 211376)
AUTHORS Hu, X., Swank, R. and Roe, B. A.
TITLE Direct Submission
JOURNAL Submitted (12-FEB-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Farrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT On Jan 29, 2002 this sequence version replaced gi:18201802.
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR
-----
FEATURES
source Location/Qualifiers
1..211376
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="3"
/clone="rp23-71el3"
/clone_lib="RPCI Mouse BAC Library 23"
BASE COUNT 60567 a 41306 c 42611 g 66892 t
ORIGIN
Query Match 7.4%; Score 51.6; DB 10; Length 211376;
Best Local Similarity 54.1%; Pred. No. 1.7;
Matches 105; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 423 AACCAATAAATAACAAAATTTCTTGTGAAGTTCCGGAGAAACTCAATCTCCACA 482
Db 52670 AAAACAACAACAACAACAACCTTTTCTCCANTGGGGGAAAAAACAACAACACCC 52611

QY 483 ATTGGATATTGGAGTATTATCGATGGATTTTCTTCTTTTAAATACATCGCGCTATA 542
Db 52610 CCAGTTTGTGCTGTPAGAATGCTTGGAATGCTACTTGTGTAATTTAGCAAGTGTGTA 52551

QY 543 AAAGAAATGCCACTTGAATAATCAAGAAATCAAAATATTCAAAATGAATAATGAAGCC 602
Db 52550 AAACCAACAAAGAACCAACCAACCAAAAAAAGAAAAAAGAAAAAAGAAAAAACA 52491

QY 603 ACTAACTCAACAA 616
Db 52490 AACAAAACAACA 52477
```


RESULT 5

PFMAL13P9 78578 bp DNA linear HTG 11-AUG-1999
LOCUS Plasmodium falciparum chromosome 13 strain 3D7, *** SEQUENCING IN
DEFINITION PROGRESS ***, in unordered pieces.

ACCESSION AL096783 GI:5731889

VERSION HTG; HTGS_PHASE1.

KEYWORDS malaria parasite P. falciparum.

SOURCE Plasmodium falciparum.

ORGANISM Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

REFERENCE 1 (bases 1 to 78578)

AUTHORS Bowman, S., Churcher, C., Harris, B., Harris, D., Lawson, D., Quail, M.

TITLE Direct Submission

JOURNAL Submitted (02-JUL-1999) P. falciparum Genome Sequencing Consortium,

The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge

COMMENT CB10 ISA, UK

On Aug 12, 1999 this sequence version replaced gi:5420293.

For more information about this sequence or the Malaria Project,

see <http://www.sanger.ac.uk/projects/P.falciparum>. IMPORTANT: This

sequence is unfinished and does not necessarily represent the

correct sequence. Work on the sequence is in progress and the

release of this data is based on the understanding that the

sequence may change as work continues. The sequence may be

contaminated with foreign sequence from E.coli, yeast, vector,

phage etc.

Order of segments is not known: 800 n's separate segments.

* NOTE: This is a 'working draft' sequence.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

FEATURES

source Location/Qualifiers

1. 78578

/organism="Plasmodium falciparum"

/strain="3D7"

/db_xref="taxon:5833"

/chromosome="13"

BASE COUNT 26762 a 9436 c 7899 g 27281 t 7200 others

ORIGIN

Query Match 7.3%; Score 51; DB 2; Length 78578;

Best Local Similarity 51.6%; Pred. No. 2.7;

Matches 142; Conservative 0; Mismatches 130; Indels 3; Gaps 1;

QY 363 AATTCTGAATATATCTTTCTTACATGGAATAACTGCATCTTGGATATATCCAAAT 422

DB 30299 AATTAAAAAATATATATTTCTGATTTTACACGGATTTTTCAGAAATATATAAT 30358

QY 423 AACACAAATAAATCAACAAATTTCTTCTGAAAGTTCGGAGAAACTCAATCTCCACA 482

DB 30359 ATGTTAGAAAAAAGGATTTTTCGATTTTAAAGAAATAAATAAATAAATAA 30418

QY 483 ATTGATATGCGATATATCGATGGATTTTCTTCTTTAAATAACATCGCGCTATA 542

DB 30419 TTTAATCATATAATATATATTTT---TTTCATGTTGATTAATAAATCTGCATCTTCTA 30475

QY 543 AAGAAATGCGCACTTGAATAATCAAGAAATCAAAATATCAAAATGAATAATGAAGCC 602

DB 30476 AAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAATATATAAATATTCATGGATA 30535

QY 603 ACTAACTCAACAAATCCAGAAATCAGGTGGGAATA 637

DB 30536 ATTGAAACTATTTATCTGGATTTCTGCTTGGATA 30570

RESULT 6

AX346425/c

LOCUS Bombyx mandarina fifth instar larva silkland mitochondrion DNA.

DEFINITION Bombyx mandarina mitochondrial DNA, complete genome.

ACCESSION AX346425

VERSION AX346425.1 GI:18494311

REFERENCE 1 (bases 1 to 15928)

AUTHORS Yukuhiro, K., Sezutsu, H., Itoh, M., Shimizu, K. and Banno, Y.

TITLE Direct Submission

KEYWORDS

SOURCE synthetic construct.

ORGANISM synthetic construct

REFERENCE 1 (sites)

AUTHORS Olek, A., Piepenbrock, C. and Berlin, K.

JOURNAL Diagnosis of diseases associated with the immune system

Patent: WO 0200928-A 1496 03-JAN-2002;

Epigenomics AG (DE)

FEATURES

source Location/Qualifiers

1. 6928

/organism="synthetic construct"

/db_xref="taxon:32630"

/note="chemically treated genomic DNA (Homo sapiens)"

BASE COUNT 1852 a 151 c 1572 g 3353 t

ORIGIN

Query Match 7.3%; Score 50.8; DB 6; Length 6928;

Best Local Similarity 47.6%; Pred. No. 4.8;

Matches 180; Conservative 0; Mismatches 197; Indels 1; Gaps 1;

QY 304 CTGAGAATAAATGACCTTTTGGTTCTGAGAAAGACTTAGCGACAACCTCTCCAATCTTCTA 363

DB 4902 CTAATACATCAATCTATTTTAAAAAATTAACAACTAAATTAATTAATTAATTAAT 4843

QY 364 ATTCGTAATATATCTTTTCTTACATGGAATAACTGCATTTTGGATATATTCCAATTA 423

DB 4842 ATCTATATACTAACATCAATTAATACCCAAACAAAAAATTTTACTTTTATAAAAACCTAAT 4783

QY 424 ACACAAATAAATCAACAAATTTCTTCTGAAAGTTCGGAGAAACCTCAATCTCCACAA 483

DB 4782 AAACATAAATACCCGATATAAATATCTATATACATAAATAAATAAATAAATAAATAA 4723

QY 484 TTGATATTTGGATATATATCATGATGATTTTCTTCTTCTTAAATAAATCAATCGCTATAA 543

DB 4722 AAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4664

QY 544 AAAAAATGCGCACTTGAATAATCAAGAAATCAAAATTAATTAATAATGAATGAAGCCA 603

DB 4663 AAAAAATCAACAACTAAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4604

QY 604 CTAACACTCAACAAATCCAGATCAGGTGGGAATAACTTAATCAAGAACCAACCAACAAAG 663

DB 4603 TACTAACTACAAATAACTTCAACACCATATATTATACCTAAAAAATAAATAAATAAATAA 4544

QY 664 CCTCATCCAGTTGTTAGA 681

DB 4543 ACCACACAAATTTCTATA 4526

RESULT 7

AB070263/c

LOCUS Bombyx mandarina mitochondrial DNA, complete genome.

DEFINITION Bombyx mandarina mitochondrial DNA, complete genome.

ACCESSION AB070263

VERSION AB070263.1 GI:18640048

KEYWORDS

SOURCE

ORGANISM

Bombyx mandarina fifth instar larva silkland mitochondrion DNA.

Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata;

Ditrysia; Bombycoidea; Bombycidae; Bombyx.

REFERENCE 1 (sites)

AUTHORS Yukuhiro, K., Sezutsu, H., Itoh, M., Shimizu, K., Nho, S. K., Zhong, B.

TITLE Not trivial level of sequence divergence and sequence

rearrangements of mitochondrial genome between the wild mulberry

silkmoth, Bombyx mandarina and its close relative, the domesticated

silkmoth Bombyx mori

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 15928)

AUTHORS Yukuhiro, K., Sezutsu, H., Itoh, M., Shimizu, K. and Banno, Y.

TITLE Direct Submission

JOURNAL

Submitted (17-AUG-2001) Kenji Yukuhiro, National Institute of
Agrobiological Sciences, Insect Genetics and Evolution Department;
Oowashii-2, Tsukuba, Ibaraki 305-8634, Japan
(E-mail: kyukis@affrc.go.jp, Tel:81-298-38-6202, Fax:81-298-38-6202)
Location/Qualifiers

FEATURES

source

1. .i5928
/organism="Bombyx mandarina"
/organelle="mitochondrion"
/db_xref="taxon:7092"
/tissue_type="silkgland"
/dev_stage="fifth instar larva"
/country="Japan: Ibaraki, Tsukuba"
243. .308
/gene="cox3"

tRNA

/note="codon recognized: GGA"
/product="tRNA-Gly"
/anticodon=(pos:273. .275,aa:Gly)
/evidence=not_experimental
312. .662
/gene="nad3"
312. .662
/gene="nad3"
/codon_start=1
/transl_table=5
/evidence=not_experimental
/product="NADH dehydrogenase subunit 3"
/protein_id="BAB84641.1"
/db_xref="GI:18640049"

gene

CDS

/translation="MMIMLITFFILSNLMMMSMILSKSFMDREKCSPECGFDP
KSMRIPFSLHFELITIFLFDVEIALIFPLMKVNMILTKTSTFFIILILIG
LYHEWNQMLNWTN"
716. .779
/gene="cox3"
/note="codon recognized: GCA"
/product="tRNA-Ala"
/anticodon=(pos:744. .746,aa:Ala)
/evidence=not_experimental
830. .893
/gene="cox3"
/note="codon recognized: CGA"
/product="tRNA-Arg"
/anticodon=(pos:859. .861,aa:Arg)
/evidence=not_experimental
895. .961
/gene="cox3"
/note="codon recognized: AAC"
/product="tRNA-Asn"
/anticodon=(pos:925. .927,aa:Asn)
/evidence=not_experimental
961. .1027
/gene="cox3"

tRNA

tRNA

tRNA

tRNA

tRNA

tRNA

tRNA

tRNA

tRNA

tRNA

tRNA

tRNA

tRNA

tRNA

tRNA

tRNA

tRNA

tRNA

tRNA

tRNA

tRNA

tRNA

tRNA

tRNA

tRNA

tRNA

tRNA

tRNA

tRNA

tRNA

tRNA

tRNA

tRNA

tRNA

tRNA

gene

CDS

gene

CDS

tRNA

tRNA

gene

CDS

/product="NADH dehydrogenase subunit 5"
/protein_id="BAB84642.1"
/db_xref="GI:18640050"
/translation="MCFNSLFFLPFSMNFFFMVYFTMNGLSVFFWEIVSFNSMSI
VMSILLDMSSLFMMFVSLISSVWIFYSKSYMSNLNRRITLLVMEFVSMVLLIIS
PNLSILGWDLGLISYCLYIYQNLKSYNAGMLTALSNGIDGMVIMLVICWVMNIG
SWITFYDMFMKNDYMEVVSFMIIIAATKSAQIPFSSWLPAPMAAPTVPVSAVHSS
TLVTAGVYLLIRFNFDLDMFFKILLFLLSITLTFMFGISANIEFDLKKIATLSLSQ
LGLMSILSGFPDIAFPALLTHSMFKALLPWCAGVITHMNDIQDIRYMGGLSLYIP
LTSCLNLSLGCIPFLSGFYSKDILEMVSPNSPMLVFIILFYLTGLTVFYTFR
LIMVYINDENLMIYLNLYDEDYIMLSMFVLLPMSLLSGSFLSFISYPMIYLPF
NLKLMVYVYLSILGSLGFWVSNMKIYSNKLMLSNLSNLCCLMMFMPLNLTYSYLSYI
YLNFSQILFKNIDMGWSEIYSGKIFNILKNYSVFYNYQMNFKILYLFSLNFMIL
FIFFI"
complement(2909. .2975)
/note="codon recognized: CAC"
/product="tRNA-His"
/anticodon=(pos:2942. .2944,aa:His)
/evidence=not_experimental
complement(3035. .4375)
/gene="nad4"
complement(3035. .4375)
/gene="nad4"
/codon_start=1
/transl_table=5
/evidence=not_experimental
/product="NADH dehydrogenase subunit 4"
/protein_id="BAB84643.1"
/db_xref="GI:18640051"
/translation="MKKILFMWIPICFMKNFWLVOMLFLMMFLFMMNLNVMC
FFNSYMPYCDIISFGLLSILWICSMIASESFLKLNFFNFVFLVIFLLMLYL
TFSVNLNLLFFFEASLITLLIIGMGYSEPIQAGMYLLFTLFGSLPLMGIIY
VNDLNTMTIYFLFNNMYLLYSIMFAELVKMPYFVHLPLPKAHEAPVSGSMI
LAGIMLGGYGLLRVMIFLOEVNLKINYIILSLVGLFISMKCFQCDVMSLIAY
SSVAHSMVIGGIMIMYWGFIGSVLMSHGLCSSGMFLANINERLHRSYINK
GMNFMPSMLWFLLLSSNMAAPPSSLNLGEISLNSLSVSNISMLLMMSIFFSA
GYSLYSYIOHGKIFQGLYSFYVGLSREYLLLLHLHFWPLNMLIKVDFTSIWF"
complement(4375. .4665)
/gene="nad4L"
complement(4375. .4665)
/gene="nad4L"
/codon_start=1
/transl_table=5
/evidence=not_experimental
/product="NADH dehydrogenase subunit 4L"
/protein_id="BAB84644.1"
/db_xref="GI:18640052"
/translation="MLIYDIWFLIYFFIGNLIFVSKNHLVLLSLEPIVLISIFF
FFVLVLMFIDYDMYMLMVFLVSVCEGSLGLSILVSMIRSHGNDYFQFSML"
4670. .4735
/gene="cox3"
/note="codon recognized: ACA"
/product="tRNA-Thr"
/anticodon=(pos:4701. .4703,aa:Thr)
complement(4735. .4802)
/note="codon recognized: CCA"
/product="tRNA-Pro"
/anticodon=(pos:4768. .4770,aa:Pro)
/evidence=not_experimental
4805. .5335
/gene="nad6"
4805. .5335
/gene="nad6"
/codon_start=1
/transl_table=5
/evidence=not_experimental
/product="NADH dehydrogenase subunit 6"
/protein_id="BAB84645.1"
/db_xref="GI:18640053"
/translation="MKMLSSMLIMMTSIMVFNHPLSMGMMLMOTMTICMISGLI
KTYWFSYILFLIGGLLVFLYVSSIASNELEFLNLMKTLIIIMMIYIMMYFI
NIKYNLIENNSNNKLTMLFNDENKINLNKYNQTSFLTNMLLIYLFVNLVAV

```
gene VKTNTFYGLRSSNK"
5391..6548
/genes="cob"
CDS 5391..6548
/genes="cob"
/codon_start=1
/transl_table=5
/evidence-not_experimental
/product="cytochrome b"
/protein_id="BAB84646.1"
/db_xref="GI:18640054"
/translators="MEFSIKRTHPIIKIINSLIDLSPSNLSSNNWNGSLALCLIM
IQITGLFLMYTANEMAFYSVNICRNVNWMIRLHANGAFFFCIIHLHGR
GIYSEFNGWYLVGIMFLMFAFMGVLPWGMOSFWMATVITNLUSAIPYLG
MLVNWGGFADNATLTRYTFHLLPFILMTIHLFLHQTGSNNPLGLNSLD
KIPHPFFFKDILKILKMLTLLINPYLLGDPNFTPANPLVPVHIQPEWYF
LFAYILRSIPNKLGGVIALIMSILIIIPFTFKKIOGQVPIIOIIFWMLMIT
ILLTWIGARVENPYITGQLLTIMYFLFPLNPIIGMYWDKILFNKKNN"
6548..6613
trna /genes="cox3"
/notes="codon recognized: UCA; codon recognized:UCN"
/product="trna-Ser"
/anticodon="(pos:6578..6580,aa:Ser)
/evidence-not_experimental
complement(6639..7583)
gene="nad1"
complement(6639..7583)
CDS /genes="nad1"
/codon_start=1
/transl_table=5
/evidence-not_experimental
/product="NADH dehydrogenase subunit 1"
/protein_id="BAB84647.1"
Query Match 7.38; Score 50.8; DB 3; Length 15928;
Best Local Similarity 48.08; Pred. No. 4.1;
Matches 145; Conservative 0; Mismatches 157; Indels 0; Gaps 0;
Qy 357 TCCTTAAATTCGAATATATATCTTTTCCTTACATGGAATACTGCATTCCTGGATATATT 416
Db 10433 TTTTATTTTAAATGAATATATAGCTTGAATATATGATTATATTTAACTTAATATTATA 10374
Qy 417 CCATTAACACAAATAAATCAACAAATTTCTCTGNAAGATTCGGGAGAAATCAATC 476
Db 10373 AGAAAAAATAAAGAGAAATATATAAATTTTAAATTTATTAATTAATCAATCAATA 10314
Qy 477 TCCCAATTTGGATATTGGAGTATTATCGATGATTTCTCTCTCTTTAATTAACATCGC 536
Db 10313 TTTAAATTTAAATATTAATTAATCAATTAATTTATTTATTTATTAATTAATCAATCA 10254
Qy 537 CCTATAAAGAAATGGCCACTTGAATTAATCAAGAAATCAAAATATTTCAAAATGAAATAT 596
Db 10253 TTGATTTATTTTAAATAAATAATTAATTTAAATTTAAATTTAAATTAATTAATTAATCA 10194
Qy 597 GAAGCCACTAACTCAACAAATCAGAAATCAGTGGGGAATTAATTAATCACTAGACCAAAA 656
Db 10193 TTAATATTTTAAATTTAAATATTAATTAATCAATCAATTAATTTATTTATTTATTAATTA 10134
Qy 657 CA 658
Db 10133 CA 10132
```

```
RESULT 8
AC068138
LOCUS AC068138 168698 bp DNA linear PRI 09-JAN-2002
DEFINITION Homo sapiens BAC clone RP11-813K12 from 2, complete sequence.
ACCESSION AC068138
VERSION AC068138.6 GI:15638958
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
REFERENCE
AUTHORS Sulston,J.E. and Waterston,R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
REFERENCE
AUTHORS Isak,A., Kozlowski,A., Doebber,A. and Dixon,R.
TITLE The sequence of Homo sapiens BAC clone RP11-813K12
JOURNAL Unpublished (2001)
REFERENCE
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (28-APR-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (18-SEP-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (09-JAN-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Sep 18, 2001 this sequence version replaced gi:15148349.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: saplensew@wustl.edu
----- Summary Statistics
-----
Center project name: H_NH0813K12
-----
NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.
This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.
MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D.
McPherson, Department of Genetics, Washington University, St. Louis
MO. For additional information about the map position of this
sequence, see http://genome.wustl.edu/gsc
SOURCE INFORMATION:
The RPCI-11 human BAC library was made from the blood of one male
donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
and coworkers at the Roswell Park Cancer Institute
(http://bacpac.med.buffalo.edu)
VECTOR: pBACe3.6
NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-314B1, 2000 bp overlap; the
clone sequenced to the right is RP11-86017. Actual start of this
clone is at base position 161288 of RP11-314B1; actual end is at
base position 168698 of RP11-813K12.
Data from AC069383 was used to finish this clone, AC068138.An
```


[illegible]

Db	301955	TATATTA	TACTTCTTATATAAATAATAATTTCTTCACAAATTTTACATTAATTTCTTTACATC	301896
Qy	371	ATATATCTTTTCCATCATGGAATAACTGCATCTCTTGGATATATTCCAAATTAACACAAA	430	
Db	301895	ATAAAAAATTAATTTATACCAATTAATATATTAATTAATTAATTAATTAATTAATTAAT	301836	
Qy	431	TAAATCAACAAAATTTCTTCTTGAAGTCCGGAGAAACCTCAATCTCCACAATTTGGATA	490	
Db	301835	TTTAAAAAATCTTTAATATTTTCAACACATATAAAAAAATAAATTAATTAATTAATTAATTAAT	301777	
Qy	491	TTGGAGTATTAATCATGATTTTCTTCTTTTAAATTAACATCGGCTATATAAAGAAAA	550	
Db	301776	TTTAATTAATTAATTTATTTATTTTCTTCTTCATATAAAAAAT---ATAATAAACTAAA	301720	
Qy	551	TGGCCACTTGAATTAATCAAGAATCAAAATTAATCAAAATGAATTAATGAAGCCACTAAACT	610	
Db	301719	CGCACACTAATTAATTTAAAAAATCAACTAAACATAATACTATACCTATATATCC	301660	
Qy	611	CAACAAATCCAGAATCAGGTGGGAATAACTTTAACTCAGAACCAAAACACAAAGCC	665	
Db	301659	CAACACTTTAAAAAATAAAACGAATAAATCATCTAAAAATCAATAATTTAAAAACC	301605	
RESULT 13				
AX344564/c				
LOCUS	AX344564	349980 bp	DNA	linear
DEFINITION	Sequence 15 from Patent WO0200932.			PAT 01-FEB-2002
ACCESSION	AX344564			
VERSION	AX344564.1	GI:18492450		
KEYWORDS	synthetic construct.			
SOURCE	synthetic construct			
ORGANISM	artificial sequence.			
REFERENCE	1 (sites)			
AUTHORS	Olek.A., Piepenbrock,C. and Berlin,K.			
TITLE	Diagnosis of known genetic parameters within the mh			
JOURNAL	Patent: WO 020932-A 15 03-JAN-2002;			
FEATURES	Epigenomics AG (DE)			
source	Location/Qualifiers			
	1..349980			
	/organism="synthetic construct"			
	/db_xref="taxon:32630"			
	/note="chemically treated genomic DNA (Homo sapiens)-Original length of seq 1: 3.673778 <223>-split as follows.-seq 01 0.000.001 TO 0.349.980-seq 02 0.300.001 649.980-seq 03 600.001 949.980-seq 04 900.001 1.249.980-seq 05 1.200.001 1.549.980-seq 06 1.500.001 1.849.980-seq 07 1.800.001 2.149.980-seq 08 2.100.001 2.449.980-seq 09 2.400.001 2.749.980-seq 10 2.700.001 3.049.980-seq 11 3.000.001 3.349.980-seq 12 3.300.001 3.649.980-seq 13 3.600.001 3.673.778 <223>-Original length of seq 2: 3.673778 <223>-split as follows.-seq 14 0.000.001 TO 0.349.980-seq 15 0.300.001 649.980-seq 16 600.001 949.980-seq 17 900.001 1.249.980-seq 18 1.200.001 1.549.980-seq 19 1.500.001 1.849.980-seq 20 1.800.001 2.149.980-seq 21 2.100.001 2.449.980-seq 22 2.400.001 2.749.980-seq 23 2.700.001 3.049.980-seq 24 3.000.001 3.349.980-seq 25 3.300.001 3.649.980-seq 26 3.600.001 3.673.778"			
BASE COUNT	99616 a 3868 c 72947 g 173548 t			1 others
ORIGIN				
Query Match	7.2%;	Score 50.2;	DB 6;	Length 349980;
Best Local Similarity	50.1%;	Pred. No. 2.8;		
Matches 178;	Conservative	0;	Mismatches 173;	Indels 4;
Gaps	2;			
Qy	311	TAAATGACATTTTGGTTCTCAGAAAGACTTACGACAACTCTCCAACTCTTAATCTGA	370	
Db	1955	TATATTA	TACTTCTATAAATAAATAATTTCTTCACAAATTTTACATTAATTTCTTTACATC	1896
Qy	371	ATATATCTTTTCCATCATGGAATAACTGCATCTCTTGGATATATTCCAAATTAACACAAA	430	

Db	1895	ATAAAAAAATAATTTATTACCAATTAATAATATATATTAACTCTTTTATAAATATATAATAT	1836
Qy	431	TAAATATCAACAAAATTTCTCTTTGAAAGTTCCGGAGAAAACTCAATCTCCACAATTTGGATA	490
Db	1835	TTTAAAAAAATCTTTAAATATTTTCAACATATATAAAAAAANAAT--AATATTTAAATTA	1777
Qy	491	TTGGAGTATTATCGATGGATTTCTCTCTCTTTTAAATTAACATCGCGCTATAAAAGAA	550
Db	1776	TTTAATTAATATTTATTTATTTCTCTCTCATATAAANAAT---ATAATAAACTATA	1720
Qy	551	TGGCAGTGTGAATATCAAGAAATCAAAATATTCAAATATGAATATGAAGCCACTAAACT	610
Db	1719	CGCACAACTAATAAATTTAAAAAATCAACTTAAACATATAAATCATCATACCTATAATCC	1660
Qy	611	CAACAAATCCAGAAATCAGTGGGATAACTTTAACTCAGAACCAAAACACAAGCC	665
Db	1659	CACACITTTAAAAACTTAAACGAATAATCATCTTAAANAATCAATAATTTAAAAACC	1605
RESULT	14		
LOCUS	AX281283/c		
DEFINITION	Sequence 25 from Patent WO0177164.	11670 bp	linear
ACCESSION	AX281283		
VERSION	AX281283.1	GI:16608538	
KEYWORDS			
SOURCE	synthetic construct.		
ORGANISM	synthetic construct		
REFERENCE	artificial sequence.		
AUTHORS	1 (sites)		
TITLE	Olek,A., Piepenbrock,C. and Berlin,K.		
JOURNAL	Diagnosis of diseases associated with apoptosis		
	Patent: WO 0177164-A 25 18-OCT-2001;		
	Epigenomics AG (DE)		
FEATURES	Location/Qualifiers		
source	1..11670		
	/organism="synthetic construct"		
	/db_xref="taxon:32630"		
BASE COUNT	3287 a 238 c 2377 g 5768 t		
ORIGIN			
Query Match	7.1%;	Score 49.6;	DB 6;
Best Local Similarity	48.9%;		Length 11670;
Matches 133;	Conservative 0;	Mismatches 139;	Indels 0;
Gaps	0;		
Qy	395	TAACTGCATCTCTGGATATATTCCAATTAACACAAATAAAATCAACAAAATTTCTCTGA	454
Db	11107	TAAATTTATTCATATCTATCTATATACCTCTCTCATATAAATCGACATATATTTTCAACA	11048
Qy	455	AAGTTCGGAGAAACCAATCTCCCAATTTGGATATTTGGAGTATTATCGATGATTTTC	514
Db	11047	CTAACTTAATAATACCCCTTTTATACAACTACGTATTTTAAACATATATACCATAAATCTTA	10988
Qy	515	TTCTCTTTAATTAACATCGCCTATATAAAGAAAATGGCACCTTGAATATCAAGAATC	574
Db	10987	ATTTAATTTTCCAAAAATCAGTAAACCATATCTCTCAITTTAAAAATAAAAAATC	10928
Qy	575	AAATATCAAAATGAATATGAAGCCACTTAACTCAACAAATCCAGAATCAGGTGGGA	634
Db	10927	TAAAAATACCTCCCTCTAATAATAAATAAAAAATAAAAAATTTAAAAATAAAAAATCAA	10868
Qy	635	ATACTTAACTCAGAACCAAAACAAAGCCT	666
Db	10867	CTAATTAACAATCAACAAACAAATAAANAAT	10836
RESULT	15		
LOCUS	AX344573/c		
DEFINITION	Sequence 24 from Patent WO0200932.	34980 bp	linear
ACCESSION	AX344573		
VERSION	AX344573.1	GI:18492459	
PAT	01-FEB-2002		

KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1 (sites)
AUTHORS Olek A., Piepenbrock, C. and Berlin, K.
TITLE Diagnosis of known genetic parameters within the mhc
JOURNAL Patent: WO 0200932-A 24 03-JAN-2002;
EpiGenomics AG (DE)
FEATURES
Location/Qualifiers
1. .349980
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)-Original length of seq 1: 3.673778 <223>-split as follows.-seq 01 0.000.001 TO 0.349.980-seq 02 0.300.001 649.980-seq 03 600.001 949.980-seq 04 900.001 1.249.980-seq 05 1.200.001 1.549.980-seq 06 1.500.001 1.849.980-seq 07 1.800.001 2.149.980-seq 08 2.100.001 2.449.980-seq 09 2.400.001 2.749.980-seq 10 2.700.001 3.049.980-seq 11 3.000.001 3.349.980-seq 12 3.300.001 3.649.980-seq 13 3.600.001 3.673.778 <223>-Original length of seq 2: 3.673778 <223>-split as follows.-seq 14 0.000.001 TO 0.349.980-seq 15 0.300.001 649.980-seq 16 600.001 949.980-seq 17 900.001 1.249.980-seq 18 1.200.001 1.549.980-seq 19 1.500.001 1.849.980-seq 20 1.800.001 2.149.980-seq 21 2.100.001 2.449.980-seq 22 2.400.001 2.749.980-seq 23 2.700.001 3.049.980-seq 24 3.000.001 3.349.980-seq 25 3.300.001 3.649.980-seq 26 3.600.001 3.673.778"
BASE COUNT 100129 a 4069 c 75030 g 170752 t
ORIGIN
Query Match 7.1%; Score 49.6; DB 6; Length 349980;
Best Local Similarity 48.6%; Pred. No. 3.5;
Matches 168; Conservative 0; Mismatches 174; Indels 4; Gaps 1;
Qy 246 TATACAAAACATGTTTGAAGAAGATGACGACATCTCTTAACCCCTTTAATCGCGCT 305
Db 262705 TAAATAATACAAATATTTAAACATACTAAACAACTACTTTATATACATCTAANAATTCAAA 262646
Qy 306 GAGATAAAATCACATTTTGGTCTCGAGAAAGACTTGGGACAACTCTCCAATCTTCTTAAT 365
Db 262645 TTTAACTAAACATCGTATATTTTTCCTAACAAATATAACAAATAATCTTCTTCCCTATA 262586
Qy 366 TCTGAATATTATCTTTTCCCTTACATGGAATACTGCATCTCTGGATATATTCCAATTAAAC 425
Db 262585 CTAAATAATCTCTTTTCCATCTTTCCAAATTTTTCCAACTAATCAAAAAATAAAAAAAA 262526
Qy 426 ACAATAAAATCACAAAAATTTCTCTTGAAGTTCCGGAGAAACTCAATCTCCACAATT 485
Db 262525 AANAATAATATAAATAAANAATAAANAATAAACCCTACCTATCAACTCTCTATATTTAATATA 262466
Qy 486 GGATATTGGAGTATTATCGATGGATTTTCTCTTCTTTTAAATTAACATCGCGCTATAAAA 545
Db 262465 AAATATTCTTAAACCAAAATATAATACTTATACCTATATATCCCAACACITTTAAAAAAC 262406
Qy 546 GAAATGCCCACTTGAATTAATCAAGAAATCAAAATATTCAAAATGAA 591
Db 262405 CAAAACGAATA----AATCACCTAAATCAAAAAATTCAAAACCAA 262364

Search completed: November 2, 2002, 01:41:03
Job time : 2972.9 secs

THIS PAGE IS BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2002, 22:12:43 ; Search time 153.968 Seconds
(without alignments)
7805.766 Million cell updates/sec

Title: us-09-881-556-3
Perfect score: 700
Sequence: 1 gaattggcagcagattttt.....accgcactctacagaaaagc 700

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802.*
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	50.8	7.3	6928	ABL33523	Human immune syste
C 2	49.4	7.1	5814	ABL33561	Human immune syste
C 3	48.8	7.0	17280	AAS46771	Tumour suppressor
C 4	48.4	6.9	7040	AAS46439	Tumour suppressor
C 5	48.2	6.9	9817	ABL33368	Human immune syste
C 6	48.2	6.9	17389	ABL33414	Human immune syste
C 7	48	6.9	10528	ABL32335	Human immune syste
C 8	47.6	6.8	605	AT31530	Human 3' apolipop
C 9	47.6	6.8	10039	ABL34043	Human immune syste

C 10	47.2	6.7	5501	24	AAS61244	Human gene regulat
C 11	47.2	6.7	9110	24	ABL34428	Human immune syste
C 12	47	6.7	6123	24	ABL32820	Human immune syste
C 13	47	6.7	8392	24	ABL33490	Human immune syste
C 14	47	6.7	17893	24	ABL33365	Human immune syste
C 15	47	6.7	40862	24	ABL34072	Human immune syste
C 16	46.6	6.7	399	22	ABL19737	Human breast cance
C 17	46.6	6.7	19734	24	ABL33933	Human immune syste
C 18	46.4	6.6	5997	24	ABL33624	Human immune syste
C 19	46.2	6.6	10957	24	ABL33110	Human immune syste
C 20	46.2	6.6	15674	24	ABL32363	Human immune syste
C 21	46.2	6.6	15674	24	ABL34477	Human metastasis a
C 22	46	6.6	6668	24	ABL33697	Human immune syste
C 23	45.8	6.5	5195	24	ABL32921	Human immune syste
C 24	45.8	6.5	8592	24	ABL33983	Human immune syste
C 25	45.8	6.5	11650	22	AAS46756	Tumour suppressor
C 26	45.6	6.5	5276	24	ABL32151	Human immune syste
C 27	45.6	6.5	6591	24	ABL33250	Human immune syste
C 28	45.6	6.5	7037	24	ABL33751	Human immune syste
C 29	45.4	6.5	693	22	AAL14804	Human breast cance
C 30	45.4	6.5	5680	24	AAS61047	Human gene regulat
C 31	45.4	6.5	6061	24	ABL32141	Human immune syste
C 32	45.2	6.5	5314	24	ABL32161	Human immune syste
C 33	45.2	6.5	6191	24	ABL33371	Human immune syste
C 34	45.2	6.5	6298	24	ABL32435	Human immune syste
C 35	45.2	6.5	11049	24	ABL32669	Human immune syste
C 36	45.2	6.5	18283	24	AAS61363	Human gene regulat
C 37	45	6.4	6031	22	AAS46621	Tumour suppressor
C 38	45	6.4	6394	24	AAS61278	Human gene regulat
C 39	45	6.4	6565	22	AAS46465	Tumour suppressor
C 40	45	6.4	7900	22	AAS46760	Tumour suppressor
C 41	45	6.4	13449	24	ABL33384	Human immune syste
C 42	45	6.4	18154	24	ABL32254	Human immune syste
C 43	44.8	6.4	5474	24	ABL33041	Human immune syste
C 44	44.8	6.4	6409	22	AAS46495	Tumour suppressor
C 45	44.8	6.4	7037	24	ABL33750	Human immune syste

ALIGNMENTS

RESULT 1
ABL33523/C
ID ABL33523 standard; DNA: 6928 BP.

XX AC ABL33523;

XX XX 26-MAR-2002 (first entry)

XX DE Human immune system associated gene SEQ ID NO: 1496.

XX KW Human; immune system disease; cytosine methylation; antiaesthatic;
KW antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmologic;
KW antirheumatic; antiarthritic; antidiabetic; antipeptidic;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.

XX OS Homo sapiens.

XX PN WO200200928-A2.

XX XX 03-JAN-2002.

XX PF 02-JUL-2001; 2001WO-EP07537.

XX PR 30-JUN-2000; 2000DE-1032529.

XX PR 01-SEP-2000; 2000DE-1043826.

XX PA (EPIG-) EPIGENOMICS AG.

OS Homo sapiens.
PN WO200168912-A2.
XX
XX
XX 20-SEP-2001.
XX
XX 15-MAR-2001; 2001WO-EP02955.
XX
XX 15-MAR-2000; 2000DE-1013847.
PR 06-APR-2000; 2000DE-1019058.
PR 07-APR-2000; 2000DE-1019173.
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX WPI; 2001-602752/68.
XX
XX Fragments of chemically modified genes associated with tumour suppressor
PT genes and oncogenes, useful in designing primers and probes for
PT analysing diseases associated with cytosine methylation state e.g.
PT cancer
XX
XX Claim 1; SEQ ID No 495; 27pp; English.
XX
XX The invention relates to a nucleic acid comprising a sequence of 18
CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
CC bisulphite, of genes associated with tumour suppression and
CC oncogenes having a sequence taken from 536 (actually 533 since
CC numbers 408, 458 and 500 are missing from the sequence listing) sequences
CC (Ss) and sequences complementary to (Ss). The nucleic acid may be a
CC peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may
CC form part of a set of probes for detecting the cytosine methylation state
CC and/or single nucleotide polymorphisms and also to be used in an
CC array for analysing diseases associated with CpG dinucleotides e.g.
CC cancers and tumours. The probes can also be used in a method for
CC ascertaining genetic and/or epigenetic parameters for the diagnosis
CC and/or therapy of existing diseases or the predisposition to specific
CC diseases, by analysing cytosine methylations. The parameters may be
CC compared to another set of genetic and/or epigenetic parameters, the
CC differences serving as basis for diagnosis and/or prognosis events which
CC are disadvantageous to patients. The present sequence is one of the
CC 533 genomic sequences derived from tumour suppressor genes and
CC oncogenes.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX
XX Query Match 7.08; Score 48.8; DB 22; Length 17280;
Best Local Similarity 47.48; Pred. No. 0.37; Mismatches 162; Indels 0; Gaps 0;
Matches 146; Conservative 0;
XX
XX 371 ATATTATCTTTCTTACATGAATACTGCATTCTTGATATATATCCAAATTAACACAAA 430
DB III III III III III III III III III III III III III III III III
XX 8458 ATAAATATTATTTTATATATATATATTTATATATAAAATTTATTTATATATATT 8399
XX
XX 431 TAAATCAACAAATTTCTTTGAAAGTTCGGGAGAAATCAATTCACCAATTTGGATA 490
DB III III III III III III III III III III III III III III III III
XX 8398 TATAACATAAAATTTTATTTATATAATATATTTATATAATAAAATTTATTTTATATA 8339
XX
XX 491 TTGAGTATTATTCATGGAATTTCTTCTTTTAAATTAACATGCGCTATAAAGAAAA 550
DB III III III III III III III III III III III III III III III III
XX 8338 TATTATATATAAAATTTTATTTTATATAATAAAATTTATTTATATAAATAATA 8279
XX
XX 551 TGGCCACTTGAATCAAGAAATCAAAATATTCAAAATGAATTAATGAAGCCACTTAAACT 610
DB III III III III III III III III III III III III III III III III
XX 8278 TTTATTTTATAATACTACTACTTAAATATAATAATAAATAACACACCACTAAAA 8219

QY 611 CACCAATCCAGATCAGTGGGAATAACTTAATCAGAACCAACACAAAGCCTCATC 670
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 8218 CTTCAAAATATATCAAAAAAATAAATTCCTATAATAACTCAAAACCAAAAAATTTAT 8159
QY 671 CAGTTGTT 678
DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 8158 TATTTT 8151
RESULT 4
AAS46439/C
ID AAS46439 standard; DNA; 7040 BP.
XX
XX AAS46439;
XX
XX 18-DEC-2001 (first entry)
XX
XX Tumour suppressor gene derived chemically modified sequence #161.
DE
XX Human; tumour suppressor gene; oncogene; antitumour; cytostatic;
KW cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
KW cytosine methylation; ds.
XX
XX Homo sapiens.
XX
XX WO200168912-A2.
PN
XX 20-SEP-2001.
XX
XX 15-MAR-2001; 2001WO-EP02955.
XX
XX 15-MAR-2000; 2000DE-1013847.
PR 06-APR-2000; 2000DE-1019058.
PR 07-APR-2000; 2000DE-1019173.
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX WPI; 2001-602752/68.
XX
XX Fragments of chemically modified genes associated with tumour suppressor
PT genes and oncogenes, useful in designing primers and probes for
PT analysing diseases associated with cytosine methylation state e.g.
PT cancer
XX
XX Claim 1; SEQ ID No 161; 27pp; English.
XX
XX The invention relates to a nucleic acid comprising a sequence of 18
CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
CC bisulphite, of genes associated with tumour suppression and
CC oncogenes having a sequence taken from 536 (actually 533 since
CC numbers 408, 458 and 500 are missing from the sequence listing) sequences
CC (Ss) and sequences complementary to (Ss). The nucleic acid may be a
CC peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may
CC form part of a set of probes for detecting the cytosine methylation state
CC and/or single nucleotide polymorphisms and also to be used in an
CC array for analysing diseases associated with CpG dinucleotides e.g.
CC cancers and tumours. The probes can also be used in a method for
CC ascertaining genetic and/or epigenetic parameters for the diagnosis
CC and/or therapy of existing diseases or the predisposition to specific
CC diseases, by analysing cytosine methylations. The parameters may be
CC compared to another set of genetic and/or epigenetic parameters, the
CC differences serving as basis for diagnosis and/or prognosis events which
CC are disadvantageous to patients. The present sequence is one of the
CC 533 genomic sequences derived from tumour suppressor genes and
CC oncogenes.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
CC

XX
SQ Sequence 7040 BP; 2018 A; 184 C; 1566 G; 3272 T; 0 other;
Query Match 6.9%; Score 48.4; DB 22; Length 7040;
Best Local Similarity 52.9%; Pred. No. 0.4;
Matches 127; Conservative 0; Mismatches 111; Indels 2; Gaps 1;
QY 423 AACACAAATAAATCAACAAATTTCTTGAAAGTTCGGAGAAAACATCTCCACA 482
DB 1830 AAACATATTTAATAATATTAATTTAAAATATCTATCAAACTAAAATRACTCG 1771
QY 483 ATGGATATGGAGTATTATCGATGATTTCTTCTTTTAATTAACATCGGCTATA 542
DB 1770 TTTTCATATAACPAACATACAAATTAATATTTTCTTCTTAAATTTTACGNATTA 1711
QY 543 AAGAAAATGGCCACTTGAATATCAAGAAATCAAAATTTCAAAATGAATGAAGCC 602
DB 1710 AAATATAAATCCTACTATATTAATTAATACACATATATAAAACGTAATAAATAAT 1651
QY 603 ACTAACTCAACAATCCAGATCAGGTGGGAATACCTTAACCTCAGAACCAACAAA 662
DB 1650 A--ACACTAATCAATCAATATATACCTATATATCATCCCCCAAAAACAAAACAAA 1593
RESULT 5
ABL33368/C
ID ABL33368 standard; DNA; 9817 BP.
XX
AC ABL33368;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 1341.
XX
KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cytotatic; neutropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antididiabetic; antipsoriatic;
KW antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200200928-A2.
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP07537.
XX
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-130909/17.
XX
PT Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation -
XX
PS Claim 1; SEQ ID NO 1341; 32pp + Sequence Listing; German.
XX
CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel

CC diseases. The present sequence is a gene of the invention.
XX
SQ Sequence 9817 BP; 2814 A; 174 C; 2033 G; 4794 T; 2 other;
Query Match 6.9%; Score 48.2; DB 24; Length 9817;
Best Local Similarity 48.4%; Pred. No. 0.46;
Matches 134; Conservative 0; Mismatches 143; Indels 0; Gaps 0;
QY 354 CAATCTCTTAATCTCGAATATTATCTTTCTTACATGGAATAAATCTCTTGATAT 413
DB 1839 CAACCTTTAAATTCCTAATTCATTAATTAATAAATAATATACATATCTACTTACT 1780
QY 414 ATTCATTAATTAACACAAATAAATCAACAAATTTCTTTGAAAGTTCGGAGAAAAC 473
DB 1779 ACTAAAATTAACACACATTAATTAATCTCTTAAAAAATTTCTACTCAAAAATTA 1720
QY 474 ATCTCCACAAATGGATATTGGAGTATTATCGATGATTTCTTCTTCTTAAATTAAC 533
DB 1719 ATTTCTAAATTAATCTCAAAACCTTAATACAAATAAATTAACATAATTAATCACT 1660
QY 534 GCGCCTATAAAGAAAATGGCCACTTGAATATCAAGAAATCAAAATATTTCAAAATGA 593
DB 1659 AAACCTTAATAACATTAATTAATCTACAGGAAAACCTAAATATATATATATATA 1600
QY 594 AATGAAGCCACTAAACTCAACAAATCCAGATCAGGT 630
DB 1599 AATAATTCATAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1563
RESULT 6
ABL33414
ID ABL33414 standard; DNA; 17389 BP.
XX
AC ABL33414;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 1387.
XX
KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cytotatic; neutropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antididiabetic; antipsoriatic;
KW antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200200928-A2.
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP07537.
XX
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-130909/17.
XX
PT Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation -
XX
PS Claim 1; SEQ ID NO 1387; 32pp + Sequence Listing; German.
XX
CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences

CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.

XX
SQ Sequence 17389 BP; 4806 A; 201 C; 3688 G; 8694 T; 0 other;

Query Match 6.9%; Score 48.2; DB 24; Length 17389;
Best Local Similarity 47.0%; Pred. No. 0.49;
Matches 149; Conservative 0; Mismatches 168; Indels 0; Gaps 0;
QY 14 GATTTTTCCTTTTACCTATTCATTTAGTTTCTTTGATTCACCAAGTC 73
DB 779 GATTTTTCCTTTTACCTATTCATTTAGTTTCTTTGATTCACCAAGTC 838
QY 74 ATTATTTGTTTAAATCCAGATGATCCGGAATTTTGAATAAATTTCTACTAAATTCGA 133
DB 839 TTTAGAGGTATGTTAAATCTTTAAATATTTAAATGAAGTTGGGTTATAGGTA 898
QY 134 AATTAATTTGACCTTTGGCTTGTAATGCTAGTTGGCTCGGAGTGATATTTTATCCT 193
DB 899 ATTGAATTTTATTTTATTTTGGTTTGGTTAGATGCTTTAGTAATATTAATATTTAT 958
QY 194 TGATAGAACACGCTCTTGAAGCTGTAGTTATTCATCGCTTGTGTTTCTATACMAA 253
DB 959 TTAATTTGGAATAATGTTTTTAAATATTTATCGATTATGTTATTTTATGATATAAT 1018
QY 254 AACATCTTTTGAAGAAGTGAAGACATTCCTCTTAACCCCTTTAATCGCGCTGAGATAA 313
DB 1019 ATTATGTTGATGATGATATTTATTTGTTATGTTAGATTTTGTATGATGAAGATAA 1078
QY 314 AATCACTTTTGGTCTG 330
DB 1079 AATGATGCTGTTTG 1095

RESULT 7
ABL32335/C
ID ABL32335 standard; DNA; 10528 BP.

XX
AC ABL32335;
XX
XX
XX 26-MAR-2002 (first entry)
XX Human immune system associated gene SEQ ID NO: 308.
XX
XX Human; immune system disease; cytosine methylation; antiasthmatic;
XX antiarteriosclerotic; anti-anaemic; cytosine methylation; antiasthmatic;
XX neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
XX antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
XX antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
XX acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
XX neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
XX gene; ds.

XX Homo sapiens.
XX
XX WO200200928-A2.
XX
XX 03-JAN-2002.
XX
XX 02-JUL-2001; 2001WO-EP07537.
XX
XX 30-JUN-2000; 2000DE-1032529.
XX 01-SEP-2000; 2000DE-1043826.
XX
XX (EPIC-) EPIGENOMICS AG.
XX Olek A, Piepenbrock C, Berlin K;
XX
XX WPI; 2002-130909/17.

XX
PT Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation -
XX
XX Claim 1; SEQ ID NO 308; 32pp + Sequence Listing; German.
XX
XX The present invention provides a number of human immune system associated
XX genes which are modified by the methylation of cytosines. The sequences
XX can be used in the diagnosis and treatment of immune system disorders,
XX including eye diseases such as retinopathy, neovascular glaucoma and
XX macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
XX leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
XX rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
XX diseases. The present sequence is a gene of the invention.

XX
SQ Sequence 10528 BP; 2873 A; 86 C; 2164 G; 5405 T; 0 other;

Query Match 6.9%; Score 48; DB 24; Length 10528;
Best Local Similarity 50.0%; Pred. No. 0.51;
Matches 146; Conservative 0; Mismatches 145; Indels 1; Gaps 1;
QY 326 TTCTGAGAAAGACTTAGCGACAACTCTCCCAATCTTCTAATTTCTCAATATTTATCTTTCT 385
DB 4679 TTCTTAAATCAATCCTTAATATTAATACTTATAATCTTATTAATAAATTTTA 4620
QY 386 TACATGGAATTAATGCTTCTTGGATATATTCATTTTAAACAAATTAACAAATTAACAAAT 445
DB 4619 TATACCTCATTTCAATTAACATTTATTAATAGCTCTTTTAAAC-CAATACTATCATATAAAC 4561
QY 446 TTCTTGAAGTTCCGGAGAAAACCTCAATCTCCCAATTTGGATATTTGGAGTATTATCGA 505
DB 4560 TAAATAATAAACAATCAATCAATCTTAAATCTTAAATAACAATCAATATATAACAATAA 4501
QY 506 TGGATTTTCTTCTTCTTAAATTAACATGCGCTTATAAAGAAAATGCCACTTTGAATAA 565
DB 4500 CATATAAATACCTTCTTAATACATACATATATATATAAATAAATAATATATACAAAA 4441
QY 566 TCAAGATCAAAATATTTCAAAATGAAATAATGAAGCCCACTAAACCTCAACAAA 617
DB 4440 ATCAATTTAACTTCACAAAAAATAAATTTAAACAAAAAATATTTAAAAA 4389

RESULT 8
AAT31530/C
ID AAT31530 standard; cDNA; 605 BP.
XX
AC AAT31530;
XX
XX 15-SEP-1996 (first entry)
XX Human 3' apolipoprotein B SAR element clone Rh32.
XX
XX Erythropoietin; EPO; anaemia; gene therapy; vector;
XX scaffold attachment region; SAR element; apolipoprotein B;
XX transgenic animal; ss.

XX Homo sapiens.
XX
XX WO9619573-A1.
XX
XX 27-JUN-1996.
XX
XX 18-DEC-1995; 95WO-CA00696.
XX
XX 19-DEC-1994; 94US-0358918.
XX
XX (CANG-) CANGENE CORP.
XX Delcuve G;
XX
XX WPI; 1996-309587/31.

XX PF 06-APR-2001; 2001WO-EP03968.
XX PR 06-APR-2000; 2000DE-1019058.
XX PR 07-APR-2000; 2000DE-1019173.
XX PR 30-JUN-2000; 2000DE-1032529.
XX PR 01-SEP-2000; 2000DE-1043826.
XX PA (EPIG-) EPIGENOMICS AG.
XX PI Olek A, Piepenbrock C, Berlin K;
XX DR WPI; 2002-017470/02.
XX PT New nucleic acid sequences from chemically modified genes associated
PT with gene regulation, useful for analysing cytosine methylations for
PT diagnosis and therapy of diseases e.g. severe combined immunodeficiency
PT disease -
XX PS Claim 1; SEQ ID No 204; 26pp; English.
XX CC The invention relates to 224 nucleic acid sequences comprising at least
CC 18 bases of a chemically pretreated gene associated with gene regulation
CC selected from 43 known genes (or complementary sequences). The
CC chemical pretreatment converts cytosine bases unmethylated at the
CC 5-position to uracil or another base with hybridisation behaviour
CC dissimilar to cytosine, to enable analysis of cytosine methylations.
CC The DNA sequences, oligomers (or sets/arrays) and method are
CC useful in the diagnosis of diseases (or predisposition to diseases)
CC associated with gene regulation and in therapy of such diseases, by
CC enabling analysis of the cytosine methylation patterns of such genes,
CC kits are provided. They are especially useful in diagnosis
CC and therapy of e.g. severe combined immunodeficiency disease, cardiac
CC disorders, haemophilia, solid tumours and cancer, Werner syndrome,
CC asthma, HDR syndrome, Saethre-Chotzen syndrome, renal disease,
CC pre-eclampsia, graft versus-host disease. The present sequence is a
CC sequence included in the sequence data for this specification and is
CC associated with the human gene regulation-associated genes.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX SO Sequence 5501 BP; 1484 A; 107 C; 1262 G; 2648 T; 0 other;
Query Match 6.7%; Score 47.2; DB 24; Length 5501;
Best Local Similarity 46.6%; Pred. No. 0.68;
Matches 151; Conservative 0; Mismatches 173; Indels 0; Gaps 0;
QY 334 AAGACTTAGGCACAACTCTCAATCTTCTTAATCTTGAATATATATCTTTCTTACATGGA 393
Db 2933 AAAATATAAAAAACAATTAATAAATAAATTAATAAATAAATAAATAAATAAATAA 2874
QY 394 ATAATCGATCTTGGATATATTCCTCAATCTTCTTAATCTTGAATATATATCTTTCTT 453
Db 2873 TTAATACTATTAACTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2814
QY 454 AAAGTTCGGAGAAACTCAATCTTCAATCTTGAATATATATCTTTCTTACATGATTTT 513
Db 2813 AAAATCTTTAAACATTTTATCTTTAAATACCTATTATAAATAAATAAATAAATAA 2754
QY 514 CTTCCTTTTAAATTAACATGGCCCTATTAAGAATAATGGCCACTTGAATATCAAGANT 573
Db 2753 ACAACCTTAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2694
QY 574 CAAATATTTCAAAATGAATTAATGAAGCCCTCAATCAATCAATCAATCAATCAATCA 633
Db 2693 ANAACTTTCAAACTAACATACCATTTTAACAATAAATAAATAAATAAATAAATAA 2634
QY 634 AATACTTAACCTAGAACCAAAAC 657
Db 2633 AATTAAAAAATAAATAAATAA 2610

RESULT 11

ABL34428/c
ID ABL34428 standard; DNA; 9110 BP.
XX AC ABL34428;
XX DT 26-MAR-2002 (first entry)
XX XX Human immune system associated gene SEQ ID NO: 2401.
XX DE Human; immune system disease: cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.
XX OS Homo sapiens.
XX PN W0200200928-A2.
XX PD 03-JAN-2002.
XX PF 02-JUL-2001; 2001WO-EP07537.
XX PR 30-JUN-2000; 2000DE-1032529.
XX PR 01-SEP-2000; 2000DE-1043826.
XX PA (EPIG-) EPIGENOMICS AG.
XX PI Olek A, Piepenbrock C, Berlin K;
XX DR WPI; 2002-130909/17.
XX PT Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation -
XX PS Claim 1; SEQ ID NO 2401; 32pp + Sequence Listing; German.
XX CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.
XX SQ Sequence 9110 BP; 2699 A; 146 C; 2072 G; 4193 T; 0 other;
Query Match 6.7%; Score 47.2; DB 24; Length 9110;
Best Local Similarity 46.6%; Pred. No. 0.73;
Matches 151; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 334 AAGACTTAGGCACAACTCTCAATCTTCTTAATCTTGAATATATATCTTTCTTACATGGA 393
Db 6542 AAAATATAAAAAACAATTAATAAATAAATAAATAAATAAATAAATAAATAAATA 6483
QY 394 ATAATCGATCTTGGATATATTCCTCAATCTTGAATATATATCTTTCTTCTTGT 453
Db 6482 TTAATACTATTAACTAAAAATAAATAAATAAATAAATAAATAAATAAATAAATA 6423
QY 454 AAAGTTCGGAGAAACTCAATCTTCAATCTTGAATATATATCTTTCTTACATGATTTT 513
Db 6422 AAAATCTTTAAACATTTTATCTTTAAATACCTATTATAAATAAATAAATAAATA 6363
QY 514 CTTCCTTTTAAATTAACATGGCCCTATTAAGAATAATGGCCACTTGAATATCAAGANT 573
Db 6362 ACAACCTTAATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 6303

Qy	574	CAAAATATTCCAAATGAATAATGAGGCACCTAAACTCAACAATTCAGAATCAGGTGGG	633
Db	6302	AAAAAATTTTCAAATAACCATACCATTTAACACAAATAAAACTAACACAAATTTAAAAAAACT	6243
Qy	634	AATAACTTAATTCAGAACCAAAAC	657
Db	6242	AATTAATAAACTAAATATATAAAC	6219
 RESULT 12 ABL32820/C			
ID	ABL32820 standard; DNA; 6123 BP.		
XX	AC		
XX	AC		
XX	DT	26-MAR-2002 (first entry)	
XX	DE	Human immune system associated gene SEQ ID NO: 793.	
XX	KW	Human; Immune system disease; cytosine methylation; antiasthmatic;	
KW	KW	antiarteriosclerotic; antianaemic; cytostatic; nootropic;	
KW	KW	nephroprotective; anti-HIV; anticongulsant; ophthalmological;	
KW	KW	antirheumatic; antiarthritic; antidiabetic; antipsoriatic;	
KW	KW	antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;	
KW	KW	acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;	
KW	KW	neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;	
KW	KW	gene; ds.	
OS	Homo sapiens.		
XX	XX		
PN	W020020928-A2.		
XX	XX		
PD	03-JAN-2002.		
XX	XX		
XX	02-JUL-2001; 2001WO-EP07537.		
XX	XX		
PR	30-JUN-2000; 2000DE-1032529.		
PP	01-SEP-2000; 2000DE-1043826.		
XX	XX	(EPIG-) EPIGENOMICS AG.	
PA	XX		
PI	Olek A, Piepenbrock C, Berlin K;		
XX	WIPI; 2002-130909/17.		
XX	XX		
PT	Nucleic acid comprising fragment of chemically modified gene, useful		
PT	for diagnosis and treatment of diseases associated with abnormal		
PT	cytosine methylation -		
XX	XX		
PS	Claim 1; SEQ ID NO 793; 32pp + Sequence Listing; German.		
XX	XX		
CC	The present invention provides a number of human immune system associa-		
CC	genes which are modified by the methylation of cytosines. The sequence		
CC	can be used in the diagnosis and treatment of immune system disorders,		
CC	including eye diseases such as retinopathy, neovascular glaucoma and		
CC	macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid		
CC	leukemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,		
CC	rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel		
CC	diseases. The present sequence is a gene of the invention.		
XX	XX		
SQ	Sequence 6123 BP; 1829 A; 102 C; 1209 G; 2983 T; 0 other;		
 Query Match 6.7%; Score 47; DB 24; Length 6123; Best Local Similarity 49.8%; Pred. No. 0.77; Matches 119; Conservative 0; Mismatches 120; Indels 0; Gaps			
Qy	419	AATTAACAAATAAAATCAACAAAATTTCTCTTGAAAGTTCCGGAGAAAACCAATCTC	478
Db	641	AATTTAAATATAAAAAACAAACAAATAAACTTCCAACAACTTCAATCTCTATCAC	582
Qy	479	CACAATGGATATGGAGPATATTCGATGGATTTCTTCCTTTAATTAACATCGGCC	538

Qy	301	CGCGTGAAGATAAAATGACTTTTGTTCTCAGAGAAGACTTTAGCGACAACTCTCCAATCTTT	360
Db	27535	ATACTAAAAATAAAAACAACAAAATACAAATTCTACTAAATATCTACTACTTTTCACACCAT	27476
Qy	361	CTAATTCGTAATATATCTTTTTCCCTCATCGGAATAACTGCATTCCTTTGGATATATATTCCA -	419
Db	27475	CGTAAAAATTAAAAAATCCTAAATCAAACCACTCACAAATCAGAAACTTATATATACTATTCTAT	27416
Qy	420	--ATTAAACACAAATAAAAATCAACAAAATTTCTCTTGAAAGTTCGGGAGAAAACCTCAATCT	477
Db	27415	TTATATAATACCAATYAACCTTAACTACTCTTAAAAATATATCAAAAAAATAAAATCAATCA	27356
Qy	478	CCACAATTGGATATGGAGTATTTATCGATGGATTTCTCTCTTTAAATTAACATCGCG	537
Db	27355	ATACCAATTAATTTACAAAAAATCTATTTAAAAAACCATCCTCAACATAAATAAAAA	27296
Qy	538	CTATAAAAAGAAAATGGCCACTTGAATTAATCAAGAAATCAAAATATTTCAAAATGAATTAATG	597
Db	27295	AATAACAAAAAACAACACCATTTATATATTTTAAACAATAAAAAAATAAAAAACTAAAA	27236
Qy	598	AAGCCACTAAACTCAACAAATCCAGAATCAGGTGGGAATTAACTTAACTCAGAACCAAAAC	657
Db	27235	ACTAAATAAAAAANAANAANAANAANAATTAACCTATCGAATACTAAATAAAAAANAANA	27176
Qy	658	ACAAGCGCTCA	668
Db	27175	AAAAAAAAACTTA	27165

Search completed: November 1, 2002, 23:23:04
Job time : 282.968 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2002, 22:17:03 ; Search time 34.9206 Seconds
(without alignments)
4923.835 Million cell updates/sec

Title: US-09-881-556-3
Perfect score: 700
Sequence: 1 gaattcgccagcagattttt.....accgcattctacagaaaagc 700

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents-NA:*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	43.6	6.2	665	2	US-08-883-795A-36
C 2	40.6	5.8	3515	2	US-08-391-743A-1
C 3	40	5.7	4163	4	US-09-004-838-70
C 4	40	5.7	4208	4	US-09-004-838-1
C 5	39.6	5.7	3706	4	US-08-913-159-9
C 6	39.4	5.6	1104	1	US-08-202-186-17
C 7	39.4	5.6	1105	1	US-08-202-186-18
C 8	39.2	5.6	176373	3	US-09-128-155-17
C 9	38	5.4	6138	4	US-09-067-800-4
C 10	38	5.4	6138	4	US-09-349-677-4
C 11	37.8	5.4	1103	1	US-08-202-186-16
C 12	37.8	5.4	1110	1	US-08-202-186-14
C 13	37.8	5.4	2030	2	US-08-705-937-7
C 14	37.4	5.3	10660	2	US-08-267-803B-8
C 15	37.4	5.3	10660	4	US-09-041-886-16
C 16	37	5.3	9636	1	US-08-323-170B-1
C 17	37	5.3	9636	4	US-08-954-441-1
C 18	36.8	5.3	1431	4	US-09-316-083-2
C 19	36.8	5.3	1539	4	US-07-956-483-21
C 20	36.8	5.3	1539	4	US-08-472-240A-13
C 21	36.8	5.3	2571	1	US-08-254-358-3
C 22	36.8	5.3	2571	1	US-08-475-391-3
C 23	36.8	5.3	2571	2	US-08-709-609-3
C 24	36.8	5.3	2571	5	PCT-US95-07178-3
C 25	36.8	5.3	3807	2	US-08-417-210A-78
C 26	36.8	5.3	4529	2	US-08-449-645A-16
C 27	36.8	5.3	4529	2	US-08-702-367A-16

C	28	36.8	5.3	4529	5	PCT-US95-04681-16	Sequence 16, Appl
	29	36.8	5.3	6474	3	US-08-651-472-66	Sequence 66, Appl
	30	36.8	5.3	6474	4	US-08-358-928-66	Sequence 66, Appl
C	31	36.8	5.3	6926	3	US-08-651-472-69	Sequence 69, Appl
C	32	36.8	5.3	6926	4	US-08-358-928-69	Sequence 69, Appl
C	33	36.8	5.3	9739	1	US-08-022-835-1	Sequence 1, Appl
C	34	36.8	5.3	9739	1	US-08-388-809-1	Sequence 1, Appl
C	35	36.8	5.3	9739	2	US-08-647-714-1	Sequence 1, Appl
C	36	36.8	5.3	9746	1	US-08-022-835-3	Sequence 3, Appl
C	37	36.8	5.3	9746	1	US-08-388-809-3	Sequence 3, Appl
C	38	36.8	5.3	9746	2	US-08-647-714-3	Sequence 3, Appl
C	39	36.6	5.2	615	4	US-08-998-416-186	Sequence 186, App
C	40	36.6	5.2	4700	2	US-08-928-692-16	Sequence 16, Appl
C	41	36.4	5.2	665	2	US-08-883-795A-36	Sequence 36, Appl
C	42	36.4	5.2	3138	1	US-07-867-106-4	Sequence 4, Appl
C	43	36.2	5.2	396	1	US-08-470-179-48	Sequence 48, Appl
C	44	36.2	5.2	982	3	US-08-973-068-28	Sequence 28, Appl
C	45	36.2	5.2	1109	1	US-08-202-186-13	Sequence 13, Appl

ALIGNMENTS

RESULT 1
US-08-883-795A-36/c
; Sequence 36, Application US/08883795A
; Patent No. 5985607
; GENERAL INFORMATION:
; APPLICANT: Delcove, Genevieve
; APPLICANT: Avang, Gregor
; TITLE OF INVENTION: Recombinant DNA Molecules and Expression
; TITLE OF INVENTION: Vectors for Tissue Plasminogen Activator
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BERESKIN & PARR
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/883,795A
; FILING DATE: 27-JUN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Gravelle, Micheline
; REGISTRATION NUMBER: 40,261
; REFERENCE/DOCKET NUMBER: 7841-062
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 665 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: Rh 32
; US-08-883-795A-36

Query Match 6.2%; Score 43.6; DB 2; Length 665;
Best Local Similarity 49.6%; Pred. No. 0.091;
Matches 112; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 383 CCTTACATGAATACTGCTTGGATATATTCCTCAATTAACACAAATAAATCAACAA 442
Db 397 CATTTTAATTAATAATAATGTAATTAACACATTTTAAATTAATAAATATGTAATTATA 338
QY 443 AATTTCTCTTGAAGTTCGGGAGAAAACCTCAATCTCCCAATTTGGATATTTGAGTATTAT 502
Db 337 CATTTTAATTAATAATAATGTAATTAACACATTTTAAATTAATAAATATTTAATTATA 278
QY 503 CGATGATTTTCTCTCTCTTAAATTAACATCGCCTATATAAGAAAATGGCCACTTGAA 562
Db 277 CATTTTAATTAATAATAATTTAATTAATAAATTTTAAATTAATAAATATTTAATTATA 218
QY 563 TAATCAAGAAATCAAAATATTCAAATGAATAATGAAGCCACTTAA 608
Db 217 TATTTTAATTAATAATAATTTAATTAATAAATATTTAATTAATAA 172
RESULT 2
US-08-391-743A-1
; Sequence 1, Application US/08391743A
; Patent No. 5843705
; GENERAL INFORMATION:
; APPLICANT: DiTullio, Paul A.; Meade, Harry; Cole, Edward S.
; TITLE OF INVENTION: TRANSGENETICALLY PRODUCED ANTI-THROMBIN III
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD, LLP
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/391.743A
; FILING DATE: 21-FEB-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Paul Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: TCI-045
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3515 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: YES
; FRAGMENT TYPE: N-terminal
US-08-391-743A-1
Query Match 5.8%; Score 40.6; DB 2; Length 3515;
Best Local Similarity 48.1%; Pred. No. 0.61;
Matches 115; Conservative 0; Mismatches 124; Indels 0; Gaps 0;
QY 307 AGAATAAAATGACTTTTGGTCTGAGAAGACTTAGCGACAACCTCTCCAACTTCTTAAT 366
Db 1404 ATAAACACTGTATTATCTTTTGGTTACATTAATTTGGCAACCCACTCCAGTACTCTTG 1463
QY 367 CTGATATTAATCTTTCTCTTACATCGAATAACTGCATTTCTTGGATATTTCCCAATTACA 426

Db 1464 ACAATATGAATTTTCTTTAAAGCTAAACCTGATTTTATTTTATTTTCCAAAGGAAT 1523
QY 427 CAATTAATAATCAACAAATTTCTCTTGAAGTTCGGGAGAAAACCTCAATCTCCAAATTTG 486
Db 1524 CTATTACACACATCAATAGGTAAACCTTCATATTAAACGTATCATTTTAAAAATTT 1583
QY 487 GATATTGGAGTATTATCGATGATTTTCTCTTCTTTAAATTAACATCGCGCTATATA 545
Db 1584 CATGTTTGAATTTTAAACAGCATTTCTTTATGTGTGATTTTATTTTACCAGAAA 1642
RESULT 3
US-09-004-838-70
; Sequence 70, Application US/09004838
; Patent No. 6350933
; GENERAL INFORMATION:
; APPLICANT: Michelmore, Richard W.
; APPLICANT: Shen, Kathy
; APPLICANT: Meyers, Blake
; TITLE OF INVENTION: Procedures and Materials for
; TITLE OF INVENTION: Conferring Pest Resistance in Plants
; NUMBER OF SEQUENCES: 140
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/004.838
; FILING DATE: 09-JAN-1998
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/781,734
; FILING DATE: 10-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Einhorn, Gregory P.
; REGISTRATION NUMBER: 38,440
; REFERENCE/DOCKET NUMBER: 023070-078810US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 70:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4163 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; FEATURE:
; NAME/KEY:
; LOCATION: 1..4163
; OTHER INFORMATION: /note= "RLG1-El69"
US-09-004-838-70
Query Match 5.7%; Score 40; DB 4; Length 4163;
Best Local Similarity 52.9%; Pred. No. 0.87;
Matches 108; Conservative 0; Mismatches 95; Indels 1; Gaps 1;
QY 393 AATAACTGCATCTTGGATATATTTCCCAATTAACAAATAAATAAATTTCTCTT 452
Db 177 AATAATGCATATTATTTTCTTTAAATAACGCATATATATATAGATTA-AAATCAT 235
QY 453 GAAAGTTCCGGAGAAAACCTCAATCTCCCAATTCGATATTCGAGTATATATCGATGGATT 512
Db 236 ATAATACATAGGTAAACTCATATATATATATATGTTTCATCCCAAGTTTATTTATATGCT 295

Db	296	CATCCTTAATTTATTATTATTATTATTATTATTATTATTATTAGAGTAGATGATCTTGTGTGATAATTAAAAAT
Qy	573	TCAAAATATTTCAAAANTGAAATAAT
Db	356	TTAATTTTGTTCAAANAATTTAAAAAT

```

RESULT 5
US-08-913-159-9/c
; Sequence 9, Application US/08913159
; Patent No. 6300109
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Plasmid-derived type II
; TITLE OF INVENTION: restriction-modification systems from Lactococcus lactis
; NUMBER OF SEQUENCES: 14
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/913.159

```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: DK 0179/95
FILING DATE: 17-FEB-1995
INFORMATION FOR SEQ ID NO: 9:

INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3706 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 ORIGINAL SOURCE:
 ORGANISM: *Lactococcus lactis* subsp. *cremoris*
 STRAIN: W56
 FEATURE:

NAME/KEY:	CDS
LOCATION:	complement (422..2161)

IDENTIFICATION METHOD: experimental
OTHER INFORMATION: /codon_start= 422

OTHER INFORMATION: /product= "LlABI methylase"
OTHER INFORMATION: /evidence= EXPERIMENTAL

OTHER INFORMATION: /gene= "ORF"
OTHER INFORMATION: /number= 1

```
OTHER INFORMATION: /standard_name=
OTHER INFORMATION: /label= m-llabi
```

FEATURE:
NAME/KEY: CDS

LOCATION: 2464..3360
IDENTIFICATION METHOD: experimental

```
OTHER INFORMATION: /codon_start= 2464
OTHER INFORMATION: /product= "LlaBI endonuclease"
```

```
Query Match      5.7%; Score 39.6; DB 4; Length 3706;
Best Local Similarity 47.9%; Pred. No. 1.1;
Matches 114; Conservative 0; Mismatches 124; Indels 0; Gaps 0;
```

Query Match 5.7%; Score 39.6; DB 4; Length 3706;
Best Local Similarity 47.0%; Pred No 1 1.

Best Local Similarity 47.36; Pseud. NO. 1.1; Matches 114; Conservative 124; Indels 0; Gaps 0

339 TTACGACAACCTCCAATCTTCTAATTCTGAATATTATCTTTTCTTACATGGAATAAC 398

679 TTACCTAAGATATTGAACCTCAGTTATTATGAATATTATATTGAAAACACTACTTATATG 620

399 TGCATTCTTGGATATATATCCAAATTAACACAAATAAAATCAACAAAATTTCTCTTGAAAGT 458

Db 619 ATAGGCGGAGTACTTTGGCTTCAAAAAAATATCTTAAAAATTTTCAATTCACAT 560
QY 459 TCGGGAACAACTCAATCTCCACAAATGGATATTGGAGTATTATCGATGGATTTTCTTCT 518
Db 559 TTGAAAGACGAATTAATCTTTATTAATAATGGATCTCAAAAGAGAGCTTGATTA 500
QY 519 TCTTTAATTAACATGCGCTATATAAGAAATGGCCACTTGAATATCAAGAAATCAA 576
Db 499 TATTTAATGAAAAATACCAATTAATAATAAATATGAAGAGTATATGAATAACGAA 442

RESULT 6

US-08-202-186-17/c

; Sequence 17, Application US/08202186

; Patent No. 5756708

; GENERAL INFORMATION:

; APPLICANT: KARAN, Mirko

; APPLICANT: BURNS, Thomas M.

; APPLICANT: DALE, James L.

; APPLICANT: HARDING, Robert M.

; TITLE OF INVENTION: DNA Sequences of Banana Bunchy Top Virus

; NUMBER OF SEQUENCES: 60

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Foley & Lardner

; STREET: 3000 K Street, N.W.

; CITY: Washington, D.C.

; COUNTRY: USA

; ZIP: 20007-5109

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/202,186

; FILING DATE: 24-FEB-1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: JEFFERY, Donald D.

; REGISTRATION NUMBER: 19,980

; REFERENCE/DOCKET NUMBER: 71611/102 FIKI

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202 672 5300

; TELEFAX: 202 672 5399

; TELEX: 904136

; INFORMATION FOR SEQ ID NO: 17:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1104 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: circular

US-08-202-186-17

Query Match

Best Local Similarity 5.6%; Score 39.4; DB 1; Length 1104;

Matches 82; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 379 TTTTCTTACATGGAATCACTGCATTTCTTGGATATATTCCAATTAACACAAATAAAATCA 438
Db 854 TATTTCCCGCTTGGATATTCATTTTAATTTCTTCTTAATAAACCATAGTTTAATAT 795
QY 439 ACAAAATTTCTTGAAGTTCCGGAGAAACTCAATCTCCACAATTTGGATATTGGAGTA 498
Db 794 TCCTCTTTGCATCTGGGAATCAAAATATACTATATCTCTCATATAATATACAATCTACAT 735
QY 499 TTATCGATGGATTTTCTTCTTTTAATTAAC 531
Db 734 ATATCCAATGATTTTCTCTCTGGCGAATAAAAC 702

RESULT 7

US-08-202-186-18/c

; Sequence 18, Application US/08202186
; Patent No. 5756708
; GENERAL INFORMATION:
; APPLICANT: KARAN, Mirko
; APPLICANT: BURNS, Thomas M.
; APPLICANT: DALE, James L.
; APPLICANT: HARDING, Robert M.
; TITLE OF INVENTION: DNA Sequences of Banana Bunchy Top Virus
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/202,186
; FILING DATE: 24-FEB-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: JEFFERY, Donald D.
; REGISTRATION NUMBER: 19,980
; REFERENCE/DOCKET NUMBER: 71611/102 FIKI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202 672 5300
; TELEFAX: 202 672 5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: circular
US-08-202-186-18

Query Match

Best Local Similarity 5.6%; Score 39.4; DB 1; Length 1105;

Matches 82; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 379 TTTTCTTACATGGAATCACTGCATTTCTTGGATATATTCCAATTAACACAAATAAAATCA 438
Db 855 TATTTCCCGCTTGGATATTCATTTTAATTTCTTCTTAATAAACCATAGTTTAATAT 796
QY 439 ACAAAATTTCTTGAAGTTCCGGAGAAACTCAATCTCCACAATTTGGATATTGGAGTA 498
Db 795 TCCTCTTTGCATCTGGGAATCAAAATATACTATATCTCTCATATAATATACAATCTACAT 736
QY 499 TTATCGATGGATTTTCTTCTTTTAATTAAC 531
Db 735 ATATCCAATGATTTTCTCTCTGGCGAATAAAAC 703

RESULT 8

US-09-128-155-17/c

; Sequence 17, Application US/09128155

; Patent No. 6117654

; GENERAL INFORMATION:

; APPLICANT: Pan, Yang

; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY

; FILE REFERENCE: 09404/052001

; CURRENT APPLICATION NUMBER: US/09/128,155

; CURRENT FILING DATE: 1998-08-03

; EARLIER APPLICATION NUMBER: US 60/091,650

; EARLIER FILING DATE: 1998-07-02

; EARLIER APPLICATION NUMBER: US 60/054,646

; EARLIER FILING DATE: 1997-08-04

; NUMBER OF SEQ ID NOS: 18


```

; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 176373
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(176373)
; OTHER INFORMATION: n = A,T,C or G
US-09-128-155-17

Query Match          5.6%; Score 39.2; DB 3; Length 176373;
Best Local Similarity 58.68; Pred. No. 2.7;
Matches 68; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 13 AGATTTTTCCTTTTACCTATTTCATTTGTTTCTTTGATTCAAAACGATGCAAGT 72
DB 87545 AGTTTATTTTTCCTTCCCATTTTCACCTAGTTGTGAAATTTTANGGACACAAAG 87486

QY 73 CATTATTTGTTTAAATCCAGATGCGGAATTTTGAAMAAACATTTCTACTAAA 128
DB 87485 CCATGAGCATATTATTCACCTCTCAGCTTTAACTGTCAATATATTTTACTAAA 87430

RESULT 9
US-09-067-800-4/c
; Sequence 4, Application US/09067800
; Patent No. 6198024
; GENERAL INFORMATION:
; APPLICANT: Yanofsky, Martin F.
; APPLICANT: Ferrandiz, Cristina
; TITLE OF INVENTION: Seed Plants Characterized by Delayed
; TITLE OF INVENTION: Seed Dispersal
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/067,800
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/067,800
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UD 2948
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6138 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..6138
; OTHER INFORMATION: /label= AGL5_promoter
; OTHER INFORMATION: /note= "Nucleotide sequence of the AGL5 promoter."
US-09-067-800-4

Query Match          5.4%; Score 38; DB 4; Length 6138;
Best Local Similarity 51.88; Pred. No. 2.7;
Matches 86; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

US-09-067-800-4

QY 13 AGATTTTTCCTTTTACCTATTTCATTTGTTTCTTTGATTCAAAACGATGCAAGT 72
DB 87545 AGTTTATTTTTCCTTCCCATTTTCACCTAGTTGTGAAATTTTANGGACACAAAG 87486

QY 73 CATTATTTGTTTAAATCCAGATGCGGAATTTTGAAMAAACATTTCTACTAAA 128
DB 87485 CCATGAGCATATTATTCACCTCTCAGCTTTAACTGTCAATATATTTTACTAAA 87430

RESULT 9
US-09-067-800-4/c
; Sequence 4, Application US/09067800
; Patent No. 6198024
; GENERAL INFORMATION:
; APPLICANT: Yanofsky, Martin F.
; APPLICANT: Ferrandiz, Cristina
; TITLE OF INVENTION: Seed Plants Characterized by Delayed
; TITLE OF INVENTION: Seed Dispersal
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/067,800
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/067,800
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UD 2948
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6138 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..6138
; OTHER INFORMATION: /label= AGL5_promoter
; OTHER INFORMATION: /note= "Nucleotide sequence of the AGL5 promoter."
US-09-067-800-4

Query Match          5.4%; Score 38; DB 4; Length 6138;
Best Local Similarity 51.88; Pred. No. 2.7;
Matches 86; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

US-09-067-800-4

QY 13 AGATTTTTCCTTTTACCTATTTCATTTGTTTCTTTGATTCAAAACGATGCAAGT 72
DB 87545 AGTTTATTTTTCCTTCCCATTTTCACCTAGTTGTGAAATTTTANGGACACAAAG 87486

QY 73 CATTATTTGTTTAAATCCAGATGCGGAATTTTGAAMAAACATTTCTACTAAA 128
DB 87485 CCATGAGCATATTATTCACCTCTCAGCTTTAACTGTCAATATATTTTACTAAA 87430

RESULT 9
US-09-067-800-4/c
; Sequence 4, Application US/09349677
; Patent No. 6286305
; GENERAL INFORMATION:
; APPLICANT: Yanofsky, Martin F.
; APPLICANT: Ferrandiz, Cristina
; TITLE OF INVENTION: Seed Plants Characterized by Delayed
; TITLE OF INVENTION: Seed Dispersal
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/349,677
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/067,800
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UD 2948
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6138 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..6138
; OTHER INFORMATION: /label= AGL5_promoter
; OTHER INFORMATION: /note= "Nucleotide sequence of the AGL5 promoter."
US-09-349-677-4

Query Match          5.4%; Score 38; DB 4; Length 6138;
Best Local Similarity 51.88; Pred. No. 2.7;
Matches 86; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

US-09-349-677-4

QY 418 CAATTAACACAAATAAATAATCAACAAAATTTCTCTTGAAGATTCGCGAGAAAACCTCAATCT 477
DB 3739 CAAATAAAAAACAATATTTTGAAGAACATTTATCTTTTAAAAAGCAGACAAAATAATCAAAAT 3680

QY 478 CCACAATTTGGATATTGGAGTATTATCGATGGATTTTCTCTCTTTTAAATAACATGCGC 537
DB 3679 GGTGAATGAACATGAGCTGTGTAATAGCTGGAAACGATATAGAAAACGGAAATGACAAAT 3620

QY 538 CTATATAAGAAAGAAATGGCCACTTGAATAATCAAGAATCAAAATATTC 583
DB 3619 CCCCAACTTGNAGTGGAGAAAATAACAAATCAAAAATCAAAATTTTC 3574

RESULT 10
US-09-349-677-4/c
; Sequence 4, Application US/09349677
; Patent No. 6286305
; GENERAL INFORMATION:
; APPLICANT: Yanofsky, Martin F.
; APPLICANT: Ferrandiz, Cristina
; TITLE OF INVENTION: Seed Plants Characterized by Delayed
; TITLE OF INVENTION: Seed Dispersal
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/349,677
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/067,800
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UD 2948
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6138 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..6138
; OTHER INFORMATION: /label= AGL5_promoter
; OTHER INFORMATION: /note= "Nucleotide sequence of the AGL5 promoter."
US-09-349-677-4

Query Match          5.4%; Score 38; DB 4; Length 6138;
Best Local Similarity 51.88; Pred. No. 2.7;
Matches 86; Conservative 0; Mismatches 80; Indels 0; Gaps 0;

US-09-349-677-4

QY 418 CAATTAACACAAATAAATAATCAACAAAATTTCTCTTGAAGATTCGCGAGAAAACCTCAATCT 477
DB 3739 CAAATAAAAAACAATATTTTGAAGAACATTTATCTTTTAAAAAGCAGACAAAATAATCAAAAT 3680

QY 478 CCACAATTTGGATATTGGAGTATTATCGATGGATTTTCTCTCTTTTAAATAACATGCGC 537
DB 3679 GGTGAATGAACATGAGCTGTGTAATAGCTGGAAACGATATAGAAAACGGAAATGACAAAT 3620
```

Db 3679 GGTGAATGACATGAGCTGTTTAAATAGCTGGAACGTATAGAAAACGGGAATGACAAGT 3620
QY 538 CTATAAAGAAATGGCCACTTGAATAATCAAGAAATCAAAATATTC 583
Db 3619 CCCCAACTTGAAGTGGAGAAATAACAATCAAAAATCAAAATTTTC 3574

RESULT 11

US-08-202-186-16/c
; Sequence 16, Application US/08202186
; Patent No. 5756708
; GENERAL INFORMATION:
; APPLICANT: KARAN, Mirko
; APPLICANT: BURNS, Thomas M.
; APPLICANT: DALE, James L.
; APPLICANT: HARDING, Robert M.
; TITLE OF INVENTION: DNA Sequences of Banana Bunchy Top Virus
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/202,186
; FILING DATE: 24-FEB-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: JEFFERY, Donald D.
; REGISTRATION NUMBER: 19,980
; REFERENCE/DOCKET NUMBER: 71611/102 FIRE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202 672 5300
; TELEFAX: 202 672 5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1103 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: circular
US-08-202-186-16

Query Match 5.4%; Score 37.8; DB 1; Length 1103;
Best Local Similarity 52.9%; Pred. No. 2.1;
Matches 81; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 379 TTTTCTTACATGGAATAACTGCATTTCTGGATATATTCCTCAATTAACACAAATAAATCA 438
Db 853 TATTTCCCGCTTGAATAATTCATTTTAAATCTCTTAAGCCATAGTTTAATAT 794
QY 439 ACAAAATTTCTCTGAAAGTCCGGAGAAACATCAATCTCCACAATTTGGAGTA 498
Db 793 TCCTCTTTGCATCTGGATATCAATAATCAATATCTCTCATATTAATACATCTACAT 734
QY 499 TTATCGATGGATTTCTTCTTTTAAATTAAC 531
Db 733 ATATCCAATGATTTTCTCTCTCGGAATAAACAAC 701

RESULT 12

US-08-202-186-14/c
; Sequence 14, Application US/08202186
; Patent No. 5756708
; GENERAL INFORMATION:
; APPLICANT: KARAN, Mirko
; APPLICANT: BURNS, Thomas M.

; APPLICANT: DALE, James L.
; APPLICANT: HARDING, Robert M.
; TITLE OF INVENTION: DNA Sequences of Banana Bunchy Top Virus
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/202,186
; FILING DATE: 24-FEB-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: JEFFERY, Donald D.
; REGISTRATION NUMBER: 19,980
; REFERENCE/DOCKET NUMBER: 71611/102 FIRE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202 672 5300
; TELEFAX: 202 672 5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1110 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: circular
US-08-202-186-14

Query Match 5.4%; Score 37.8; DB 1; Length 1110;
Best Local Similarity 52.9%; Pred. No. 2.1;
Matches 81; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 379 TTTTCTTACATGGAATAACTGCATTTCTGGATATATTCCTCAATTAACACAAATAAATCA 438
Db 852 TATTTCCCGCTTGAATAATTCATTTTAAATCTCTTAATTAACCAATTAATTAATAA 793
QY 439 ACAAAATTTCTCTGAAAGTCCGGAGAAACATCAATCTCCACAATTTGGAGTA 498
Db 792 TCCTCTTTGCATCTAGGATATCAATAATCAATATCTCTCATATTAATACATCTACAT 733
QY 499 TTATCGATGGATTTCTTCTTTTAAATTAAC 531
Db 732 ATATCCAATGATTTTCTCTCTCGGAATAAACAAC 700

RESULT 13

US-08-705-937-7
; Sequence 7, Application US/08705937
; Patent No. 5981841
; GENERAL INFORMATION:
; APPLICANT: Santino, Colleen G.
; APPLICANT: Conner, Timothy W.
; TITLE OF INVENTION: EARLY SEED 5' REGULATORY SEQUENCE
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carmen Rodriguez, Paralegal, Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30

ADDRESSEE: Campbell & Flores LLP
STREET: 4370 La Jolla Village Drive, Suite 700
CITY: San Diego
STATE: California
COUNTRY: United States
ZIP: 92122

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/041,886
FILING DATE:

CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 2626
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 10660 base pairs
TYPE: nucleic acid
SUBSTANCE:

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2002, 23:27:34 ; Search time 1235.45 Seconds
(without alignments)
7647.311 Million cell updates/sec

Title: US-09-881-556-3

Perfect score: 700

Sequence: 1 gaattggcagcagattttt.....accgcattctacagaaagc 700

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estnu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_htc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_htc:*
- 12: gb_gss:*
- 13: gb_gss_hum:*
- 14: em_gss_inv:*
- 15: em_gss_pln:*
- 16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	64.2	9.2	1101	12	CNS0039G
2	61	8.7	1092	12	CNS020K7
C 3	55.2	7.9	1190	12	CNS020N7
4	54.8	7.8	836	12	CNS01100
C 5	54.8	7.8	928	12	CNS00DKY
C 6	54.8	7.8	1101	12	CNS0100X
C 7	54.6	7.8	1101	12	CNS000B8
C 8	54.6	7.8	1101	12	CNS00EVL
C 9	53.8	7.7	1101	12	CNS002FG
C 10	53.8	7.7	1101	12	CNS0145U
C 11	53.2	7.6	1101	12	CNS0039G
12	52.4	7.5	1101	12	CNS00EVL
13	52.4	7.5	1101	12	CNS017KX
14	52.2	7.5	1043	12	CNS0145P
C 15	52.2	7.5	1200	12	CNS016CO
16	51.8	7.4	1101	12	CNS003BD
17	51.6	7.4	867	10	BE782580

C 18	51.6	7.4	1101	12	CNS00PYG
C 19	51.4	7.3	987	12	CNS00K3Q
20	51.2	7.3	759	12	CNS007BZ
C 21	51.2	7.3	1101	12	CNS00FEV
C 22	51.2	7.3	1200	12	CNS007EV
23	51	7.3	759	12	CNS060XV
24	51	7.3	1101	12	CNS003DX
C 25	51	7.3	1201	12	CNS0167M
26	50.6	7.2	1091	12	CNS014AC
C 27	50.6	7.2	1101	12	CNS00FMC
C 28	50.4	7.2	805	12	CNS06V7I
29	50.4	7.2	1101	12	CNS012TP
30	50.2	7.2	886	12	BH177277
31	50.2	7.2	886	12	CNS07JUX
32	50.2	7.2	1101	12	CNS016LI
33	49.8	7.1	664	12	BH185157
34	49.8	7.1	664	12	CNS07PXL
35	49.8	7.1	928	12	CNS00DKY
36	49.8	7.1	931	12	CNS005IS
37	49.6	7.1	1176	12	CNS03R83
C 38	49.4	7.1	1101	12	CNS003BD
C 39	49	7.0	949	12	CNS04AII
C 40	49	7.0	1029	12	CNS012GM
C 41	48.8	7.0	919	12	CNS005RL
C 42	48.8	7.0	945	12	CNS04DOK
43	48.8	7.0	1101	12	CNS0039Q
44	48.6	6.9	306	9	AU037189
C 45	48.6	6.9	379	9	AL514359

ALIGNMENTS

RESULT 1	CNS0039G	1101 bp	DNA	linear	GSS 03-JUN-1999
CNS0039G	Drosophila melanogaster genome survey sequence TET3 end of BAC #				
LOCUS	BACR08K10 of RPCI-98 library from drosophila melanogaster (fruit fly), genomic survey sequence.				
DEFINITION	AL063921				
ACCESSION	AL063921.1	GI:4941778			
VERSION	GSS.				
KEYWORDS	fruit fly.				
SOURCE	Drosophila melanogaster				
ORGANISM	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE	1 (bases 1 to 1101)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage ; BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr				
COMMENT	- Web : www.genoscope.cns.fr) Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mamoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's pi and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .				
FEATURES	Location/Qualifiers				
source	1. .1101				
	/organism="Drosophila melanogaster"				
	/db_xref="taxon:7227"				
	/clone_lib="RPCI-98"				

```
/clone="BACR08K10"
/notes="end : TET3"
BASE COUNT 201 a 64 c 131 g 202 t 503 others
ORIGIN

Query Match 9.2%; Score 64.2; DB 12; Length 1101;
Best Local Similarity 17.6%; Pred. No. 0.51;
Matches 100; Conservative 255; Mismatches 212; Indels 2; Gaps 2;

QY 19 TTTTTCCTTTTACCTATTCAATAGTTCTTTTCATTCACCAAGCATGCAAAAGTCATTAT 78
Db 483 TTTTTCCTTTTACCTATTCAATAGTTCTTTTCATTCACCAAGCATGCAAAAGTCATTAT 78
QY 79 TTTTTCCTTTTACCTATTCAATAGTTCTTTTCATTCACCAAGCATGCAAAAGTCATTAT 138
Db 543 TTTTTCCTTTTACCTATTCAATAGTTCTTTTCATTCACCAAGCATGCAAAAGTCATTAT 602
QY 139 AATTCAGCTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 198
Db 603 AATTCAGCTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 661
QY 199 GGAACACGCTCTTGAAGCTGTAAGTATTCAATCGCTTGTGTTTCTATACAAAACAT 258
Db 662 AATTCAGCTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 721
QY 259 GTTTTGAAGAATGAAGACATGCTTAAACCCCTTTTAAATCGCGCTGAGATAAATGA 318
Db 722 YTMWYMAAYMYCTA-CTYHHHHHHYHAYHTTWAWAHAMMMHHHAYAAAAAWA 780
QY 319 CTTTGTGCTGAGAAAGCTAGGACACATCTCAATCTTCAATCTTCAATCTTCAATCTT 378
Db 781 WATTHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH 840
QY 379 TTTTTCCTTACATGGAATACTGCATCTTGGATATATTCATTAACACAAAATAAATCA 438
Db 841 HHTTTHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH 900
QY 439 ACAAAATTTCTTGAAGTCCGGGAAACATCAATCTCCACATTTGGATATGAGATG 498
Db 901 HHTTTHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH 960
QY 499 TTTTTCCTTACATGGAATACTGCATCTTGGATATATTCATTAACACAAAATAAATCA 558
Db 961 HHTTTHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH 1020
QY 559 TGAATAATCAAGATCAAAATATTCAAA 587
Db 1021 ATWWMMHHWATWWMMWWWWATAAWA 1049

RESULT 2
CNS020K7 1092 bp DNA linear GSS 12-MAY-2000
LOCUS Tetraodon nigroviridis genome survey sequence T7 end of clone
DEFINITION 22211 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION ALI75696
VERSION ALI75696.1 GI:7813753
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis.

REFERENCE 1 (bases 1 to 1092)
AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bertot,A., Fizames,C., Wincker,P., Brothier,P., Quetier,F.,
Saurin,W. and Weissbach,J.
Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence
Unpublished
Genoscope.
Direct Submission
Submitted (12-APR-2000) to the EMBL/GenBank/DDAJ databases
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/tetraodon.
FEATURES
Location/Qualifiers
source 1..1092
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="22211"
/clone_lib="G"
/notes="Genoscope sequence ID : COAG222CF06LPI-end : T7"
BASE COUNT 383 a 169 c 165 g 262 t 113 others
ORIGIN

Query Match 8.7%; Score 61; DB 12; Length 1092;
Best Local Similarity 36.8%; Pred. No. 1.6;
Matches 157; Conservative 54; Mismatches 216; Indels 0; Gaps 0;

QY 236 TTGTGTTTCTATACAAAACATGTTTCAAAAAGATGAGACACATGCTTTAAACCCCTT 295
Db 645 WTTTTCCTTAAWAAAAAAWAAAAWAAAAWAAAAWAAAAWAAAAWAAAAWAAAAWAAAA 704
QY 296 TAATCGCGCTGAGATAAATGACTTTTGGTCTGAGAAGACTTAGCGACAACTCTCCA 355
Db 705 TAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 764
QY 356 ATCTTCTTAATTCGATATATATCTTTTCTTACATGGAATAACTGCAATCTTGGATAT 415
Db 765 WAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 824
QY 416 TCCAAATTAACACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 475
Db 825 WTTATWAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 884
QY 476 CTCACAAATTTGGATATGAGTATATGATGATTTCTTCTTCTTCTTCTTCTTCTTCT 535
Db 885 ATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 944
QY 536 GCCTATAAAGAAATGCGCACTTGAATAATCAAGAAATCAAAATATTCAAAATGAATAA 595
Db 945 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1004
QY 596 TGAAGCCACTAACTCAACAAATCCAGATCAGGTGGGATTAACCTTAACCTAGAACCAA 655
Db 1005 AAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1064
QY 656 ACACAAA 662
Db 1065 ATAAAAA 1071

RESULT 3
CNS020N7 1190 bp DNA linear GSS 14-MAY-2000
LOCUS Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
DEFINITION 153P04 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION AL206908
VERSION AL206908.1 GI:7865727
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis.

REFERENCE 1 (bases 1 to 1092)
AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bertot,A. and
Weissenbach,J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
Unpublished
Genoscope.
Direct Submission
Submitted (12-APR-2000) to the EMBL/GenBank/DDAJ databases
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/tetraodon.
FEATURES
Location/Qualifiers
source 1..1092
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="22211"
/clone_lib="G"
/notes="Genoscope sequence ID : COAG222CF06LPI-end : T7"
BASE COUNT 383 a 169 c 165 g 262 t 113 others
ORIGIN
```

Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
1 (bases 1 to 1190)
Roest-Crolius H., Jaillon O., Dasilva C., Fizames C., Fisher C.,
Bouneau L., Billault A., Quetier F., Saurin W., Bernot A. and
Weissenbach J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
Unpublished
2 (bases 1 to 1190)
Roest-Crolius H., Jaillon O., Dasilva C., Bouneau L., Fisher C.,
Bernot A., Fizames C., Wincker P., Brottier P., Quetier F.,
Saurin W. and Weissenbach J.
Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence
Unpublished
3 (bases 1 to 1190)
Genoscope.
Direct Submission
Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
<http://www.genoscope.cns.fr/tetraodon>.

FEATURES
source
 1..1190
 /organism="Tetraodon nigroviridis"
 /db_xref="taxon:99883"
 /clone="153P04"
 /clone_lib="G"
 /note="Genoscope sequence ID : COAGI53DH02SP1-end ;
 PUC-Ori"

BASE COUNT 373 a 229 c 141 g 312 t 135 others

Query Match 7.9%; Score 55.2; DB 12; Length 1190;
Best Local Similarity 35.3%; Pred. No. 12;
Matches 206; Conservative 64; Mismatches 313; Indels 1; Gaps 1;

QY 20 TTTTTCCTTTACCTATTTCAAATAGTTCTTGATCCAACGATGCNAAGTCATTATT 79
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1001 TATTATTGTGWKTATATATATNTATTTTTATATAWATNATRTGMTATATTWTAT 942
QY 80 TGTTTTTAATCCAGTCGGATCGGAATTTTCAAAAACATTTCTACTAAATTCGAAATTAA 139
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 941 ATATWTATATATATATATATATATATATATATATATATATATATATATATATATAT 882
QY 140 ATTTCAGCTTGGCTTGATAAATGGTAGTTGGCTCGGAGGTGATATTTTATCCTTGATAG 199
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 881 ATMTTTTAATAAANATATTTTTTATAATATWTATAWAATAWAWATATATAWAAAANAATA - 823
QY 200 GAACACGCTCTCGAAGCTGTAACTTATTCATCGCTTGCTTTTCTATACAAAACAATG 259
 ::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 822 WMMWTATATTATAAANAANAATWTAWATANTATTATATATATATATATATAMWMTTWANAATAT 763
QY 260 TTTTGAAGAAGATGAAGCACATCTGCTTTAAACCCTTTTAATCGCGTGAAGATAAAATGAC 319
 ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 762 ATATNAANAANAANAATAATATATATTTTTATATATATATATATATATATATATATATA 703
QY 320 TTTTGGTTCGAGAAAGACTTTAGCGACAACCTCCCAATCTCTTAATCTGAATATATCT 379
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 702 TTWWATATTTATATAATATTTTTTTTATATATATATATATATATATATATTTTWTWTAAATTT 643
QY 380 TTTCCTTACATGGAACTAACCTGCATCTCTGGATATATATCCCAATTAACACAATAAATCAA 439
 ||| ::| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 642 TTTTAATWTATATTATATAWATTTTWTTTTWTWMAATWMAATANATATAWAAANNATAWA 583
QY 440 CAAAATTTCTCTCGAAGCTTCGGAGAAAACCTCAATCTCCACAATTTGGATATTTGGAGTAT 499
 ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 582 ANWAATAAAWATTTAAATTTTTAAANAANAATTTTTTAATATTTTNNAAATNTTATTTAA 523
QY 500 TATCGATGGATTTTCTCTCTTTAAATTTAAACATGCGCTATAAAGAAAAATGGCCACTT 559

[illegible]

```
Db 418 AAAATTTTAAATTTATATATAAAAGTTTATATATATATATAAAGAA-TAAATTTTCAATAT 476
QY 542 AAAAGAAATGCGCCCTTGAATAATCAAGAATCAAAATATTCAAAATGAATGAAGC 601
:||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
Db 477 WAAAAGAAATGTTAAAAAATAATTTWAAAAAATAATTAATAATATATTTATTTAAAAAT 536
QY 602 CACTAAATCAACAATCCAGATCAGTGGGAATA 637
:||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
Db 537 WMTTATATATATAAATAATTAATAAWTANTNNNNAA 572

RESULT 5
CNS00DKY 928 bp DNA linear GSS 04-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC #
DEFINITION BACR27A24 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL071865.1 GI:4948170
VERSION AL071865
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 928)
REFERENCE Genoscope.
AUTHORS Direct Submission
TITLE Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammosier in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES
source
Location/Qualifiers
1..928
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RPCI-98"
/clone="BACR27A24"
/notes="end : T7"
BASE COUNT 262 a 70 c 84 g 321 t 191 others
ORIGIN
Query Match 7.8%; Score 54.8; DB 12; Length 928;
Best Local Similarity 28.9%; Pred. No. 15;
Matches 92; Conservative 83; Mismatches 143; Indels 0; Gaps 0;

QY 307 AGAATAAATGACATTTGGTTCTCAGAAAGACTTAGCGACAACCTCTCAATCTCTCAATT 366
||||| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:|
Db 875 AAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 816
QY 367 CTGAATATTATCTTTCTCTACATGGAATAACTGCATCTTGGATATATTCCAATTAAACA 426
:|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:|
Db 815 HHHHTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 756
QY 427 CAATAAATCAACAATAATCTCTTGAAGAGTTCGAGAGAAATCAATCTCCACAATTG 486
:|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:|
Db 755 NNCWAWATWWWWATATTTATATWATAWAAAAAANAANAANAANAANAANAANAANAANAANA 696
QY 487 GATATTGGAGTATATCGATGATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 546
```

```
Db 695 TTTATTTTAAAAATTTTTTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 636
QY 547 AAAATGGCCACTTGAATAATCAAGAATCAAAATATTCAAAATGAATGAAGCCTACTA 606
:||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
Db 635 TATTATATWATATTTTWWTATTTWATATATAAAATWTTATTTTAAWTAATAAATTTAA 576
QY 607 AACTCAACAATCCAGAA 624
:||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
Db 575 AAWTAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 558

RESULT 6
CNS0100X/c 1101 bp DNA linear GSS 26-JUL-1999
LOCUS Drosophila melanogaster genome survey sequence SP6 end of BAC
DEFINITION BACN03G04 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL098379
VERSION AL098379.1 GI:5609990
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
REFERENCE Genoscope.
AUTHORS Direct Submission
TITLE Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.
FEATURES
source
Location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/plasmid="pBelobAC11"
/db_xref="taxon:7227"
/clone_lib="DrosBAC"
/clone="BACN03G04"
/notes="end : SP6"
BASE COUNT 195 a 108 c 131 g 161 t 506 others
ORIGIN
Query Match 7.8%; Score 54.8; DB 12; Length 1101;
Best Local Similarity 19.1%; Pred. No. 14;
Matches 97; Conservative 195; Mismatches 217; Indels 0; Gaps 0;

QY 190 TCCTTGATAGGAACACGCTCTTGAAGCTGTAGCTATTCAATCGCTGTGTTTCTATA 249
:|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:|
Db 898 TMAATMYMYMHMYCYMYCYTMYTCTCYCHYCCYCYMYMHMATMTMCATATATMT 839
QY 250 CAAAAACATGTTTGAAGAAGATGAAGCACATTCCTTAACCCCTTTAATCGCCCTGAGA 309
:|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:|
Db 838 WATWTACAHHTTWWTAMDHMMTMMCTYHMCWYAHHHYHYCTWMMHMTTHHHCHMH 779
QY 310 ATAAATGACTTTTGGTTCTGAGAAAGACTTAGCGACAACCTCTCAATCTCTTAATCTG 369
:|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:|
Db 778 CHYTMHHHHHYHMYTAAAMMAAANAANAANAANAANAANAANAANAANAANAANAANAANA 719
QY 370 AATATATCTTTCTTACATGGAATAACTGCATCTTGGATATATTCCAATTAACACA 429
:|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:|
Db 718 THTTTTATTTTWTYVTCMTSMHCHCCACACMHYCHYCHYMHMHMHMHMHMHMHMHMHMHMH 659
QY 430 ATAAATCAACAATTTCTTCTTCAAAAGTTCGAGAGAAACTCAATCTCCCAATTCGAT 489
:|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:| :|:|:|
```


[illegible]

```

RESULT 15      CNS016CO          1200 bp   DNA       linear    GSS 26-JUL-1999
LOCUS         Drosophila melanogaster genome survey sequence T7 and of BAC
DEFINITION     BACN15E04 of DrosBAC library from Drosophila melanogaster (fruit
                fly), genomic survey sequence.
ACCESSION      ALI06578
VERSION        ALI06578.1 GI:5622626
KEYWORDS       GSS.
SOURCE         fruit fly.
ORGANISM       Drosophila melanogaster
               Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
               Prterygota; Neoptera; Endopterygota; Diptera; Brachycera;
               Muscomorpha; Ephydroidea; Drosophilidae; Drosophilla.
REFERENCE      1 (bases 1 to 1200)
AUTHORS        Genoscope.
TITLE          Direct Submission
JOURNAL        Submitted (23-JULY-1999) Genoscope - Centre National de Sequençage :
BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
              - Web : www.genoscope.cns.fr)
COMMENT        Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pheloBACL1.

FEATURES             Location/Qualifiers
     source            1..1200
                       /organism="Drosophila melanogaster"
                       /plasmid="pBelobACL1"
                       /db_xref="taxon:7227"
                       /clone_lib="DrosBAC"
                       /clone="BACN15E04"
                       /note="end :T7"
BASE COUNT           351 a   188 c   167 g   307 t   187 others
ORIGIN

Query Match                  7.5%   Score 52.2; DB 12; Length 1200;
Best Local Similarity       35.9%; Pred.No. 33;
Matches                      98; Mismatches 125; Indels 0; Gaps 0;

Qy  15 ATTTTTCCTTTTACCATTCAATAGTTCTTGTGATTCAAACGATGCNAAGTCA 74
      :|||:|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::
Db  1081 WTTTTTTTTTTWTWATWWTTTWTWTWTAATWTTTWAAATWTTATTTTATAATTWA 1022

Qy  75 TTATTGTTTAAATCCAGATCGGAATTTTGAAAACAATTTCTACTAATTCGAA 134
      :|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::
Db  1021 ATAATTTWATWTTWTTATWTTATWTTAAAAAAAWTTTTTWWTTTAAANTTWAARAW 962

Qy  135 ATAAATTTGAGCVTGGCTTGATAAATCGTAGTTTGGCTCGGAGCTGATTTTTATCCTT 194
      |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::
Db  961 ATTAAATTTAAAAAWATTTTWRTRTTWMAATTAATTTTAAAAATTTTCGTTAWAAT 902

Qy  195 GATAGGAACACGCCCTTGAAGCTGTAGTTATTCAATCGCTGCTGTTCTATACAAA 254
      :|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::
Db  901 AWTCHTTATTTTWTARAANAATTTTAAATWTTTWTWTWTTTWTWTTWWWAAACCCAAA 842

Qy  255 ACATGTTTGGAAAAGAATGAAGCACATTGTCCTT 287
      |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::
Db  841 AACTTTWAGHAARGCYYYAMCCCCCKKKKKK 809

```

Search completed: November 2, 2002, 01:47:13
Job time : 1251.45 secs

Search completed: November 2, 2002, 01:47:13
Job time : 1251.45 secs


```
source 1..1310
/organism="Cryptosporidium parvum"
/db_xref="taxon:5807"
BASE COUNT 470 a 253 c 206 g 381 t
ORIGIN
Query Match 99.0%; Score 617; DB 3; Length 1310;
Best Local Similarity 100.0%; Pred. No. 1.8e-113;
Matches 617; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCTCAATGGTGAACATCAAGATCTGGTTCAGAGCAAGCCCTATTACTCACCAGAAA 60
Db 694 CCTCAATGGTGAACATCAAGATCTGGTTCAGAGCAAGCCCTATTACTCACCAGAAA 753
QY 61 ACGAATCAAGTTCAAAATCATCTTCTGTGACAGATCCAGATCTGGATCAGTTCAAAATCT 120
Db 754 ACGAATCAAGTTCAAAATCATCTTCTGTGACAGATCCAGATCTGGATCAGTTCAAAATCT 813
QY 121 CTTCTGTTACTATTCAGAGACTGGATCAGATCAGATCAGCGCGCTTGTGACAATTC 180
Db 814 CTTCTGTTACTATTCAGAGACTGGATCAGATCAGATCAGCGCGCTTGTGACAATTC 873
QY 181 AGAGACTGGATCAGTTCAAAATCATCTTCTGTGACAGATCCAGATCTGGATCAGTTCA 240
Db 874 AGAGACTGGATCAGTTCAAAATCATCTTCTGTGACAGATCCAGATCTGGATCAGTTCA 933
QY 241 ATCACACTCTGCTACTTCTCCAGAGAGGATTTGGACTCAGAACGTTACCAATCACTTCT 300
Db 934 ATCACACTCTGCTACTTCTCCAGAGAGGATTTGGACTCAGAACGTTACCAATCACTTCT 993
QY 301 ACAGAACAACTCAAGCCAGCTACATATCTTAACCAAGAAATGAAATCATATAATC 360
Db 994 ACAGAACAACTCAAGCCAGCTACATATCTTAACCAAGAAATGAAATCATATAATC 1053
QY 361 AGAAGGTAATTCGAGTTTAAATACACTAAATCTTCCAAATCAACCCCAATCTTTCAGCA 420
Db 1054 AGAAGGTAATTCGAGTTTAAATACACTAAATCTTCCAAATCAACCCCAATCTTTCAGCA 1113
QY 421 AGCTGCAGATGTGGAAGTTATGGGAAAGAGATAAAATGGTTGATGGTGAGCAAGTAA 480
Db 1114 AGCTGCAGATGTGGAAGTTATGGGAAAGAGATAAAATGGTTGATGGTGAGCAAGTAA 1173
QY 481 TCACATAAAATGACATATTGAGATACHTTCGAAGAAATTTAGAACAAATGTAAGTA 540
Db 1174 TCACATAAAATGACATATTGAGATACHTTCGAAGAAATTTAGAACAAATGTAAGTA 1233
QY 541 TCTGCATTGATAAATATGCTTATGCAATTTCCAAATATCTAAATGTGCACTCAAGTAA 600
Db 1234 TCTGCATTGATAAATATGCTTATGCAATTTCCAAATATCTAAATGTGCACTCAAGTAA 1293
QY 601 AAAAAAAAAAAAAA 617
Db 1294 AAAAAAAAAAAAAA 1310
```

```
RESULT 2
PFMAL13P2_0
WPCOMMENT
Sequence split into 4 fragments LOCUS PFMAL13P2 Accession AL049185
Fragment Name Begin End
PFMAL13P2_0 1 110000
PFMAL13P2_1 100001 210000
PFMAL13P2_2 200001 310000
PFMAL13P2_3 300001 377835
LOCUS PFMAL13P2 377835 bp DNA linear HTG 11-AUG-1999
DEFINITION Plasmodium falciparum chromosome 13 strain 3D7, *** SEQUENCING IN
PROGRESS ***, in unordered pieces.
ACCESSION AL049185
VERSION AL049185.4 GI:5731883
KEYWORDS HTG; HTGS_PHASE1.
SOURCE malaria parasite P. falciparum.
ORGANISM Plasmodium falciparum
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
```

```
REFERENCE 1 (bases 1 to 377835)
AUTHORS Bowman,S., Churcher,C., Harris,B., Harris,D., Lawson,D., Quail,M.
and Barrell,B.
TITLE Direct Submission
JOURNAL Submitted (15-MAR-1999) P.falciparum Genome Sequencing Consortium,
The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge
CB10 1SA, UK
COMMENT
On Aug 12, 1999 this sequence version replaced gi:5531337.
For more information about this sequence or the Malaria Project,
see http://www.sanger.ac.uk/Projects/P_falciparum. IMPORTANT: This
sequence is unfinished and does not necessarily represent the
correct sequence. Work on the sequence is in progress and the
release of this data is based on the understanding that the
sequence may change as work continues. The sequence may be
contaminated with foreign sequence from E.coli, yeast, vector,
phage etc.
Order of segments is not known; 800 n's separate segments.
* NOTE: This is a 'working draft' sequence.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
FEATURES
Location/Qualifiers
source 1..377835
/organism="Plasmodium falciparum"
/strain="3D7"
/db_xref="taxon:5833"
/chromosome="13"
BASE COUNT 13224 a 36545 c 35558 g 130270 t 43238 others
ORIGIN
Query Match 9.3%; Score 58; DB 2; Length 110000;
Best Local Similarity 50.4%; Pred. No. 0.02;
Matches 142; Conservative 0; Mismatches 140; Indels 0; Gaps 0;
QY 335 CCAAGAAATGAAATCATATAATCAGAGAGGTAATTCGAGTTTATACACTAAATCT 394
Db 24936 CCACCTTATGAAATCATATAATGAAAGAGTGAATTTTAAATGAAATCAATATCA 24995
QY 395 TCCAATCAACCCCAATCTTTCACGCAAGCTGCGACATGGAAGCTTATGGGAAAGGA 454
Db 24996 AAAAAAAAAAGAAATATATATAAATAAAGAAATATGATAATAAATAAAGAA 25055
QY 455 TAAATGTTGATGGTGAGCAAGTAAATCACTAAAAATGACATTAATGAGATCTCGAA 514
Db 25056 AAATAATAATAAATAAAGAAATATAATAATAAATAAAGAAATATAATAA 25115
QY 515 AGAATTTAGAACAAATGTAAGTATCTGCATTCGATAAATATGCCCTTACCATTTCCA 574
Db 25116 TAAAAAAGAAATAATAATAATAAATAAAGAAATATAATAAATAAATAAAGAA 25175
QY 575 AATATCTAAATTTGCACTCAAGTAAAAAATAAATAAATAAATAAATAAATAA 616
Db 25176 AAATAATAATAAATAAAGAAATATAATAATAAATAAATAAATAAATAAATAA 25217
RESULT 3
AC007926/c
LOCUS AC007926 155204 bp DNA linear HTG 17-JUL-2001
DEFINITION Trypanosoma brucei chromosome II clone RPC193-3H15, *** SEQUENCING
IN PROGRESS ***, 1 ordered pieces.
ACCESSION AC007926
VERSION AC007926.8 GI:14787210
KEYWORDS HTG; HTGS_PHASE2.
SOURCE Trypanosoma brucei.
ORGANISM Trypanosoma brucei
Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
Trypanosoma.
REFERENCE 1 (bases 1 to 155204)
AUTHORS El-Sayed,N.M., Ghedin,E., Song,J., Larkin,C., Wanless,D., Jones,K.,
Peterson,J., Hou,L., Zhao,H., Mason,T., Millscher,J., Pai,G., Van
Aken,S., Utterback,T., Khalak,H.G., Gerard,C., Leech,V., Ullu,E.,
Melville,S., White,O., Adams,M.D., Donelson,J.E. and Fraser,C.M.
Trypanosoma brucei GUTat10.1 RPC193-3H15 BAC genomic sequence
```


JOURNAL Unpublished
REFERENCE 2 (bases 1 to 155204)
AUTHORS El-Sayed,N.M., Khalak,H. and Adams,M.D.
TITLE Direct Submission
JOURNAL Submitted (28-JUN-1999) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA
COMMENT On Jul 17, 2001 this sequence version replaced gi:12746529.
* NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 155204: contig of 155204 bp in length.
FEATURES
Location/Qualifiers
source
1..155204
/organism="Trypanosoma brucei"
/isolate="GUTat10.1"
/db_xref="taxon:5691"
/chromosome="II"
/clone="RPC193-3H15"
BASE COUNT 48747 a 31603 c 32329 g 42525 t
ORIGIN
Query Match 9.2%; Score 57.4; DB 2; Length 155204;
Best Local Similarity 47.6%; Pred. No. 0.025;
Matches 169; Conservative 0; Mismatches 186; Indels 0; Gaps 0;
Qy 262 AGAAGAGGATTGGACTCAGAACGTTACCAATCAGTCTTACAGAACAACTCAAGCCAG 321
Db 77407 ATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAAT 77348
Qy 322 CTACATATCCCTAACCAAGAAATGAAATCATATATATATATATATATATATATATAT 381
Db 77347 ATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 77288
Qy 382 ATACACTAAATCTTCCAAATCAAGCCAACTTTTCACGCAAGCTGGCAGATGTGGAAGTT 441
Db 77287 ATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 77228
Qy 442 ATGGGAAAGCATAAATGTTGTGTGAGCAAGTAAATCACTTAAATGACATTTATG 501
Db 77227 ATGATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 77168
Qy 502 AAGATACCTCCAAAGAAATGAAACAAATGTAAGTATCTGCATTCATTAATATGCCC 561
Db 77167 GGATAGTCTGTGAGTGTATATACGAATATATATATATATATATATATATATAT 77108
Qy 562 TTAGCCATTTCCAAATATCTAANTTGTCACTCAAGTAAAAAAGAAAAAAGAAAAA 616
Db 77107 ATTGAATTTATCAGAAAGAAATGATACAAATGAAAGAAAGAAAGATCAATCAAA 77053
RESULT 4
MGU36923 1993 bp DNA linear PLN 13-DEC-2001
LOCUS Magnaporthe grisea host-species specificity (Pw1) gene, complete
DEFINITION
ACCESSION U36923
VERSION U36923.1 GI:1045530
KEYWORDS
SOURCE Magnaporthe grisea.
ORGANISM Magnaporthe grisea
Eukaryote; Fungi; Ascomycota; Perizomycotina; Sordariomycetes;
Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.
REFERENCE 1 (bases 1 to 1993)
AUTHORS Kang,S., Sweigard,J.A. and Valent,B.
TITLE The Pw1 host specificity gene family in the blast fungus Magnaporthe grisea
JOURNAL Mol. Plant Microbe Interact. 8 (5), 939-948 (1995)

MEDLINE 96112748
PUBMED 8664503
REFERENCE 2 (bases 1 to 1993)
AUTHORS Kang,S., Sweigard,J.A. and Valent,B.
TITLE Direct Submission
JOURNAL Submitted (22-SEP-1995) Seogchan Kang, Biology, University of New Mexico, Albuquerque, NM 87131, USA
COMMENT Sweigard, J.A. Plant Cell 7, 1221-1233, 1995.
FEATURES
Location/Qualifiers
source
1..1993
/organism="Magnaporthe grisea"
/db_xref="taxon:148305"
gene
43..1215
/gene="Pw1"
CDS
join(43..111,502..558,772..1215)
/gene="Pw1"
/note="Pw1 host species specificity gene"
/codon_start=1
/product="Pw1 protein"
/protein_id="AAA80239.1"
/db_xref="GI:1045531"
/translation="MLFTALVNLPAVRVRRIDSGAPMISFNIALLLIHLHYCRFNKK
FNKTIPLYLAFSTAVIAGRKWKVYNDKRGREGSISIRKGAEGDNCGPGVPG
GPDMMRVVHEDNGNIIRGMPGKYLIGDDKEDKDNQYYSRNGYHVGDPAYQNHGGG
OMGDGYGPPGQITNQHGRQGGCHM"
BASE COUNT 778 a 313 c 372 g 530 t
ORIGIN
Query Match 9.2%; Score 57.2; DB 8; Length 1993;
Best Local Similarity 49.0%; Pred. No. 0.062;
Matches 152; Conservative 0; Mismatches 158; Indels 0; Gaps 0;
Qy 308 AAATCTCAAGCCAGCATATCTCTAACCAAGAAATGAAATCATATATATATATATAT 367
Db 1479 AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1538
Qy 368 TAATTCGAGTTTATACACATAATCTTCCAAATCAACCCCAATCTTTCACCGCAAGCTGGC 427
Db 1539 AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1598
Qy 428 AGATGTGMAAGTTATGGGAAAAGGATATAAATGTTGTGAGCAAGTAAATCACTAA 487
Db 1599 AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1658
Qy 488 AAATGACATTTATGAAGATCTTCGAAGAAATTTAGAACAAATGTAAAGTATCTGCAT 547
Db 1659 AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1718
Qy 548 TGATAAATATGCGCTTAGCCATTTCCCAATATCTAAATTTGCAACTCAAGTAAAAA 607
Db 1719 AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1778
Qy 608 AAAAAAANA 617
Db 1779 AAAAAAGAA 1788
RESULT 5
AC016068 126038 bp DNA linear HTG 13-JUL-2000
LOCUS Homo sapiens chromosome 4 clone RP11-431B5 map 4, LOW-PASS SEQUENCE
DEFINITION
ACCESSION AC016068
VERSION AC016068.3 GI:9126605
KEYWORDS HTG; HTGS_PHASE0.
SOURCE human.
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 126038)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 4, clone RP11-431B5
JOURNAL Unpublished

REFERENCE
AUTHORS

2 (bases 1 to 126038)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
 Baldwin,J., Barna,N., Beckerly,R., Boguslavskiy,L., Boukhgalter,B.,
 Brown,A., Castelle,A., Cangelino,M., Collins,S., Collamore,A.,
 Cooke,P., DeAtellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
 Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
 Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
 Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
 Lehoczy,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
 McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
 Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
 Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
 Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
 Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
 Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (20-NOV-1999) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Jul 13, 2000 this sequence version replaced gi:6715901.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

TITLE
JOURNAL
COMMENT

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L4278
 Center clone name: 431_B_5

* NOTE: This record contains 130 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

1
 * 890 989: contig of 889 bp in length
 * 990 1894: contig of 905 bp in length
 * 1895 1994: gap of 100 bp
 * 1995 2876: contig of 882 bp in length
 * 2877 2976: gap of 100 bp
 * 2977 3868: contig of 892 bp in length
 * 3869 3968: gap of 100 bp
 * 3969 4868: contig of 900 bp in length
 * 4869 4968: gap of 100 bp
 * 4969 5882: contig of 914 bp in length
 * 5883 5982: gap of 100 bp
 * 5983 6879: contig of 897 bp in length
 * 6880 6979: gap of 100 bp
 * 6980 7908: contig of 929 bp in length
 * 7909 8008: gap of 100 bp
 * 8009 8908: contig of 900 bp in length
 * 8909 9008: gap of 100 bp
 * 9009 9870: contig of 862 bp in length
 * 9871 9970: gap of 100 bp
 * 9971 10879: contig of 909 bp in length
 * 10880 10979: gap of 100 bp
 * 10980 11863: contig of 884 bp in length
 * 11864 11963: gap of 100 bp
 * 11964 12838: contig of 875 bp in length
 * 12839 12938: gap of 100 bp
 * 12939 13829: contig of 891 bp in length
 * 13830 13929: gap of 100 bp
 * 13930 14854: contig of 925 bp in length
 * 14855 14954: gap of 100 bp
 * 14955 15863: contig of 909 bp in length

* 15864 15963: gap of 100 bp
 * 15964 16850: contig of 887 bp in length
 * 16851 16950: gap of 100 bp
 * 16951 17871: contig of 921 bp in length
 * 17872 17971: gap of 100 bp
 * 17972 18858: contig of 887 bp in length
 * 18859 18958: gap of 100 bp
 * 18959 19690: contig of 732 bp in length
 * 19691 19790: gap of 100 bp
 * 19791 20687: contig of 897 bp in length
 * 20688 20787: gap of 100 bp
 * 20788 21671: contig of 884 bp in length
 * 21672 21771: gap of 100 bp
 * 21772 22673: contig of 902 bp in length
 * 22674 22773: gap of 100 bp
 * 22774 23665: contig of 892 bp in length
 * 23666 23765: gap of 100 bp
 * 23766 24662: contig of 897 bp in length
 * 24663 24762: gap of 100 bp
 * 24763 25690: contig of 928 bp in length
 * 25691 25790: gap of 100 bp
 * 25791 26696: contig of 906 bp in length
 * 26697 26796: gap of 100 bp
 * 26797 27710: contig of 914 bp in length
 * 27711 27810: gap of 100 bp
 * 27811 28693: contig of 883 bp in length
 * 28694 28793: gap of 100 bp
 * 28794 29700: contig of 907 bp in length
 * 29701 29800: gap of 100 bp
 * 29801 30699: contig of 899 bp in length
 * 30700 30799: gap of 100 bp
 * 30800 31699: contig of 900 bp in length
 * 31700 31799: gap of 100 bp
 * 31800 32718: contig of 919 bp in length
 * 32719 32818: gap of 100 bp
 * 32819 33744: contig of 926 bp in length
 * 33745 33844: gap of 100 bp
 * 33845 34734: contig of 890 bp in length
 * 34735 34834: gap of 100 bp
 * 34835 35771: contig of 937 bp in length
 * 35772 35871: gap of 100 bp
 * 35872 36789: contig of 918 bp in length
 * 36790 36889: gap of 100 bp
 * 36890 37797: contig of 908 bp in length
 * 37798 37897: gap of 100 bp
 * 37898 38819: contig of 922 bp in length
 * 38820 38919: gap of 100 bp
 * 38920 39854: contig of 935 bp in length
 * 39855 39954: gap of 100 bp
 * 39955 40860: contig of 906 bp in length
 * 40861 40960: gap of 100 bp
 * 40961 41848: contig of 888 bp in length
 * 41849 41948: gap of 100 bp
 * 41949 42837: contig of 889 bp in length
 * 42838 42937: gap of 100 bp
 * 42938 43842: contig of 905 bp in length
 * 43843 43942: gap of 100 bp
 * 43943 44850: contig of 908 bp in length
 * 44851 44950: gap of 100 bp
 * 44951 45883: contig of 933 bp in length
 * 45884 45983: gap of 100 bp
 * 45984 46879: contig of 896 bp in length
 * 46880 46979: gap of 100 bp
 * 46980 47858: contig of 879 bp in length
 * 47859 47958: gap of 100 bp
 * 47959 48839: contig of 881 bp in length
 * 48840 48939: gap of 100 bp
 * 48940 49851: contig of 912 bp in length
 * 49852 49951: gap of 100 bp
 * 49952 50886: contig of 935 bp in length
 * 50887 50985: gap of 100 bp
 * 50987 51916: contig of 930 bp in length
 * 51917 52016: gap of 100 bp

```
* 52017 52912: contig of 896 bp in length
* 52913 53012: gap of 100 bp
* 53013 53911: contig of 899 bp in length
* 53912 54011: gap of 100 bp
* 54012 54890: contig of 879 bp in length
* 54891 54990: gap of 100 bp
* 54991 55881: contig of 891 bp in length
* 55882 55981: gap of 100 bp
* 55982 56858: contig of 877 bp in length
* 56859 56958: gap of 100 bp
* 56959 57860: contig of 902 bp in length
* 57861 57960: gap of 100 bp
* 57961 58859: contig of 899 bp in length
* 58860 58959: gap of 100 bp
* 58960 59855: contig of 896 bp in length
* 59856 59955: gap of 100 bp
* 59956 60872: contig of 917 bp in length
* 60873 60972: gap of 100 bp
* 60973 61865: contig of 893 bp in length
* 61866 61965: gap of 100 bp
* 61966 62867: contig of 902 bp in length
* 62868 62967: gap of 100 bp
* 62968 63861: contig of 894 bp in length
* 63862 63961: gap of 100 bp
* 63962 64858: contig of 897 bp in length
* 64859 64958: gap of 100 bp
* 64959 65851: contig of 893 bp in length
* 65852 65951: gap of 100 bp
* 65952 66849: contig of 898 bp in length
* 66850 66949: gap of 100 bp
* 66950 67838: contig of 889 bp in length
* 67839 67938: gap of 100 bp
* 67939 68873: contig of 935 bp in length
* 68874 68973: gap of 100 bp
* 68974 69863: contig of 910 bp in length
* 69864 69963: gap of 100 bp
* 69964 70890: contig of 907 bp in length
* 70891 70990: gap of 100 bp
* 70991 71931: contig of 940 bp in length
* 71931 72030: gap of 100 bp

Query Match      8.8%  Score 54.8:  DB 2:  Length 126038;
Best Local Similarity 38.5%:  Pred. No. 0.085;
Matches 137:  Conservative 0:  Mismatches 219:  Indels 0:  Gaps 0:

QY 262 AGAAGAAGGATTGGACTCAGAACGTTTACCAATCACTTCTACAGAACAACTCAAAGCCAG 321
Db 122722 ANNAANNAANNAANNAANNAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 122781
QY 322 CTACATATCTTACCAGAAATGAAATCATATATATATATATATATATATATATATATAT 381
Db 122782 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 122841
QY 382 ATACACTAAATCTTCCAATCAACCCCAATCTTTCACGCAAGCTGGCAGATGTCGAAAGTT 441
Db 122842 ANNAANNAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 122901
QY 442 ATCGCGAAAGGATTAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 501
Db 122902 ANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA 122961
QY 502 AAGATATCTCGAAGAAATTTAGAACCAAAATGTTAAAGTATCTGCAATGATTAATATGCC 561
Db 122962 ANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA 123021
QY 562 TTACCCATCTCCCAATATCTTAATTTGTTCACTCAAGTAAANNAANNAANNAANNA 617
Db 123022 ANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA 123077

RESULT 6
AC044851
LOCUS
```

DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Homo sapiens chromosome 2 clone RP11-382G9 map 2, LOW-PASS SEQUENCE SAMPLING.
AC044851
VERSION
GI:7543818
HTG: HTGS_PHASE0.
human.

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

Homo sapiens
Eukaryota; Metazoa; Chordata; Cranial; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 70332)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Homo sapiens chromosome 2, clone RP11-382G9
Unpublished
2 (bases 1 to 70332)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bede, F., Boguslavsky, L., Bouckhgalter, B., Brown, A., Burkett, G., Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S., Domino, M., Doyle, M., Ferrel, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., LaRocque, K., Lamazara, R., Landers, T., Lehotzky, J., Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheters, R., Melidrim, J., Meneu, L., Mihova, T., Miranda, C., Mieng, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tessa, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT

Direct Submission
Submitted (12-APR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L9217
Center clone name: 382G_9

NOTE: This record contains 81 individual
sequencing reads that have not been assembled into
contigs. Runs of N are used to separate the reads
and the order in which they appear is completely
arbitrary. Low-pass sequence sampling is useful for
identifying clones that may be gene-rich and allows
overlap relationships among clones to be deduced.
However, it should not be assumed that this clone
will be sequenced to completion. In the event that
the record is updated, the accession number will
be preserved.

1 666: contig of 666 bp in length
667 766: gap of 100 bp
767 1534: contig of 768 bp in length
1535 1634: gap of 100 bp
1635 2392: contig of 758 bp in length
2393 2492: gap of 100 bp
2493 3249: contig of 757 bp in length
3250 3349: gap of 100 bp
3350 4123: contig of 774 bp in length
4124 4223: gap of 100 bp
4224 4982: contig of 759 bp in length
4983 5082: gap of 100 bp

RESULT 10

AC079427
LOCUS
DEFINITION
AC079427 235183 bp DNA linear HTG 01-SEP-2000
SEQUENCE, 50 unordered pieces.
AC079427
VERSION
AC079427.1 GI:958039
KEYWORDS
HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 235183)
DOE Joint Genome Institute.
AUTHORS
Sequencing of Mouse
TITLE
Unpublished
REFERENCE
2 (bases 1 to 235183)
DOE Joint Genome Institute.
AUTHORS
Direct Submission
TITLE
Submitted (01-SEP-2000) Production Sequencing Facility, DOE Joint
JOURNAL
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov

Project Information
Center Project Name: 0
Center clone name: RPCI-23_121D10

Summary Statistics
Consensus quality: 155133 bases at least Q40
Consensus quality: 177576 bases at least Q30
Consensus quality: 186200 bases at least Q20
Estimated insert size: 193300; agarose-fp estimation
Estimated insert size: 230283; sum-of-contigs estimation
Quality coverage: 3.62 in Q20 bases; agarose-fp estimation
Quality coverage: 3.04 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 50 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1
1408: contig of 1408 bp in length
1409: gap of unknown length
1509: contig of 1542 bp in length
3051: gap of unknown length
3151: contig of 1142 bp in length
4292: gap of unknown length
4392: gap of 1197 bp in length
5589: contig of 1197 bp in length
5689: gap of unknown length
5690: contig of 1144 bp in length
6834: gap of unknown length
6934: contig of 1480 bp in length
8414: gap of unknown length
8514: contig of 1229 bp in length
9742: gap of unknown length
9842: contig of 1157 bp in length
9843: gap of unknown length
10999: contig of 1279 bp in length
11000: gap of unknown length
12379: contig of 1125 bp in length
12479: gap of unknown length
13603: contig of 1125 bp in length
13703: gap of unknown length
13604: contig of 1847 bp in length
13704: gap of unknown length
15551: gap of unknown length
15650: contig of 1739 bp in length
15651: gap of unknown length
17390: contig of 1196 bp in length
17489: gap of unknown length
17490: contig of 1196 bp in length
18685: gap of unknown length
18686: contig of 1442 bp in length
20227: contig of 1442 bp in length
18786: gap of unknown length

FEATURES

Location/Qualifiers

20228
20328
20328
21739: contig of 1412 bp in length
21740: gap of unknown length
21840: contig of 1337 bp in length
23177: gap of unknown length
23177
23277: contig of 1106 bp in length
24382: contig of 1106 bp in length
24383
24482: gap of unknown length
24483
25663
25663
25762: gap of unknown length
25762
27496: contig of 1734 bp in length
27497
27597: gap of unknown length
27597
28446: contig of 1250 bp in length
28447
28947: gap of unknown length
30077: contig of 1131 bp in length
30078
30177: gap of unknown length
30178
31320: contig of 1143 bp in length
31321
31420: gap of unknown length
31421
32565: contig of 1145 bp in length
32566
32666: gap of unknown length
34104: contig of 1439 bp in length
34105
34204: gap of unknown length
34205
35227: contig of 1023 bp in length
35228
35327: gap of unknown length
35328
36342: contig of 1015 bp in length
36343
36442: gap of unknown length
36443
37499: contig of 1057 bp in length
37500
37599: gap of unknown length
37600
38863: contig of 1264 bp in length
38864
38963: gap of unknown length
38964
40930: contig of 1967 bp in length
40931
41030: gap of unknown length
41031
42733: contig of 1703 bp in length
42734
42833: gap of unknown length
42834
46616: contig of 3783 bp in length
46617
46716: gap of unknown length
46717
49735: contig of 3019 bp in length
49736
49835: gap of unknown length
49836
54170: contig of 4335 bp in length
54171
54270: gap of unknown length
54271
57482: contig of 3212 bp in length
57483
57582: gap of unknown length
57583
62907: contig of 5325 bp in length
62908
63007: gap of unknown length
63008
67244: contig of 4237 bp in length
67245
67344: gap of unknown length
67345
70854: contig of 3510 bp in length
70855
70954: gap of unknown length
70955
79528: contig of 8574 bp in length
79529
79629: gap of unknown length
79629
85646: contig of 6018 bp in length
85647
85746: gap of unknown length
91135: contig of 5389 bp in length
91136
91235: gap of unknown length
91236
100241: contig of 9006 bp in length
100242
100341: gap of unknown length
100342
108469: contig of 8128 bp in length
108470
108569: gap of unknown length
108570
121543: contig of 12974 bp in length
121544
121643: gap of unknown length
121644
132991: contig of 11348 bp in length
132992
133091: gap of unknown length
133092
146298: contig of 13206 bp in length
146299
146397: gap of unknown length
146398
161967: contig of 15570 bp in length
161968
162067: gap of unknown length
162068
176910: contig of 14843 bp in length
176911
177010: gap of unknown length
177011
192703: contig of 15693 bp in length
192704
192803: gap of unknown length
192804
213079: contig of 20276 bp in length
213080
213179: gap of unknown length
213180
235183: contig of 22004 bp in length.


```
source 1. 235183
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="16"
/clone="RP23-121D10"
/clone_1lb="rPCI mouse BAC library 23"
BASE COUNT 71346 a 43034 c 50925 g 59494 t 10384 others
ORIGIN

Query Match 8.5%; Score 53; DB 2; Length 235183;
Best Local Similarity 39.7%; Pred. No. 0.17;
Matches 140; Conservative 0; Mismatches 213; Indels 0; Gaps 0;

QY 265 AGAAGGATTGGACTCAGAACGTTTACCACATCTTCTACAGAACAACTCAAGCCAGCTA 324
Db 13942 AAAGGGGNNNAAAAAAAAAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 14001

QY 325 CATATCTCAACCAAGAAATGAAATCATATAATCATAGGAGGTAATTCGAGTTTAAATA 384
Db 14002 NNNNNNNNNNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA 14061

QY 385 CACTAAATCTTCCAAATCAACCCCAATCTTTTCCACCAAGCTGGCAGATGTGGAAGTATG 444
Db 14062 GAACAAAAAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA 14121

QY 445 GGGAAAGGATAAATGGTTGATCGTGAGCAAGTAATCACTAAATAATGACATTTATGAAG 504
Db 14122 AAAAAAGGAAGAGAAGAAANNGGAAAAAANNAANNAANNAANNAANNAANNAANNAAGG 14181

QY 505 ATACTTCGAAGAAATAGAACCAAAATGTAACATATCTGATTCGATTAATAATATGCGCTTA 564
Db 14182 AGAGCAAGAAAGGAAGAAAGAAAGAAAGAAAGAAAGAAAGGAGGAAAAAANNA 14241

QY 565 GCCATTTCCAAATATCTAATTTGTCACACTCAAGTAAAAAANNAANNAANNA 617
Db 14242 AGAAGAGAAAAAANAGAAAGAAAGAAAGAAAGAAAGAAAGGAGAAAGAAAA 14294

RESULT 11
AF466309 5772 bp DNA linear INV 05-FEB-2002
LOCUS Dictyostellium discoideum serine protease/ABC transporter TagD
DEFINITION (tagD) gene, complete cds.
ACCESSION AF466309
VERSION AF466309.1 GI:18496823
KEYWORDS Dictyostellium discoideum.
SOURCE Dictyostellium discoideum.
ORGANISM Eukaryota; Mycetozoa; Dictyostellida; Dictyostellium.
REFERENCE 1 (bases 1 to 5772)
AUTHORS Anjard,C. and Loomis,W.F.
TITLE Evolution of the ABC transporters of Dictyostellium
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 5772)
AUTHORS Anjard,C. and Loomis,W.F.
TITLE Direct Submission
JOURNAL Submitted (08-JAN-2002) Biology Dept 0368, UCSD, 9500 Gilman Drive, La Jolla, CA 92093, USA

FEATURES
source 1. 5772
/organism="Dictyostellium discoideum"
/strain="Ax4"
/db_xref="taxon:44689"
<243. .>5720
/gene="tagD"
/product="serine protease/ABC transporter TagD"
<243. .>5720
/gene="tagD"
243. .5720
/gene="tagD"
/notes="similar to TagB and TagC"
/codon_start=1
/product="serine protease/ABC transporter TagD"
```

```
/protein_id="AAL74253.1"
/db_xref="GI:18496824"
/translation="MKSNTNIRVLLVSLGILIFILGIRKFEFINKNNNDKIGINRKLK
FYFTKNNNNNNNNKQDLKNDKRIKLLKNEIITDNIKKIKNNKNEEIEFP
NFYSRLKSNDDMEIOQFYRKSHYIVQFKDHIINDETRQFKOFLINTDVLDEQVFP
SHIVNIPHDSFLVLANDEQSNLLSKENWSIGEFEPESNKIHLNNEKISIGLIPVLIK
LSDTNSLQORWENTLNSILTSYNSKVKLTLNQKLLKSIYVCHDRSSSSSSGLSVSS
EKUYOWISEQSESYIERSEKFTANRLSPKAITGKDTLVNNDRIDIDIPRGQGLL
SIADTGLDGHCEFFSDSNPIPYNVSNLNRKVTYVIGSLHDNEDYVDGHTIVCGSA
AGAPEDSLAISFSGLATDAKIAFFDLASDPSPNPPEPDIYOLYQPLYNAGARVII
GDSWGLSTQGLYSGDDAGSIDDFLYTHPDFILIRAGNNEQYSSLLSQATKAVNI
TVCAEOTTTHESYTTDALEYSNFETVAKSTLNSLCSQFDDKCYCTTAAQCCTEYSTVGK
WLVESGVPKPSRWYIGELGKDKKSNKESYLSSTQNVSCFYCTYKPSGSGNSG
GIPRIATLVTPDPSPYSGAKLNVNLDLTMNTSESEFIFYSSGSGSYVGTGKGTLL
GIVISVATVKTWPPSYSGAKLNVNLDLTMNTSESEFIFYSSGSGSYVGTGKGTLL
LVFTFTLSKLSLAFATVPVLAIVTQVOSKFTKRLTRLLFHNSKASOHGQSMVNM
HVSNTCKQDREIAKYSBQLMVFQISRRLLINNTFAASIKMLVLSLATIILIFGAY
LAIQKFTVGLVLSFSLYIGYVDSSTTLFGVSYSSVQCLASATRVLIILURSAPRKR
TLSEEDLDNIIDTNDNNNNNNDDISDSSDDDDNNNNNNKNNKNNKNNKNNKNNKNN
EDAEYKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNNKNN
EIDDDGDDDDGDDGDDGDDGDDGDDGDDGDDGDDGDDGDDGDDGDDGDDGDDGDD
GISCLENLIPSAYTELTCRGELEKNNVSPCYPSRADVGLVNLIDLKFSGICYGLV
CPSGSKSTLLELISRYSLHPSGGKIIYMDGIDIIAKIRPSNLASPTVNVHQHPLFLA
TIISNIGYALDNPQEDIIIEAALANAHEFIQSLPKOYDTMLTGGNLISGGKKRIAY
ARAICAKRKIMLDEITAEIDPESEAIKSIKVLTRGHTVYVVAHKVAAVRDCDKTF
VLDKGQIVEQTHNQLMMAKKYRYFAFSEDDDDYAPLLV"
BASE COUNT 2189 a 708 c 904 g 1971 t
ORIGIN

Query Match 8.4%; Score 52.6; DB 3; Length 5772;
Best Local Similarity 51.5%; Pred. No. 0.42;
Matches 121; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 303 AGAACAACTCAAGCCAGCTACATATCTTCAACCAAGAAATGAAATCAATAATAATCAG 362
Db 4605 ACAAACTATCAAAATTCACCACCATTTAGTAGTGATGATGATGATGATGATGAT 4664

QY 363 GAAGTAAATTCGAGTTTAAATACATACATAATCTTCAACCAATCAACCAATCTTTCACGCAAG 422
Db 4665 AATAATGATAATAATAATAATGATGATAATAATCAGCAAGATCCAAATAATAATAAT 4724

QY 423 CTGCAGATGTGGAAGTTATGGGGAAGAAAGGATAAAATGGTTGATGGTGGAGCAAGTAATC 482
Db 4725 GAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4784

QY 483 ACTAAATATGACATTATTGAAGATACITTCGAAAGAAATAGAAACAAATCTTAA 537
Db 4785 AATAATAATAATAATAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4839

RESULT 12
AC006281 199551 bp DNA linear HTG 12-AUG-2000
LOCUS Plasmodium falci-parum chromosome 12 clone 3D7, *** SEQUENCING IN
DEFINITION PROGRESS ***, 2 unordered pieces.
ACCESSION AC006281
VERSION AC006281.8 GI:9797738
KEYWORDS HTG; HTGS_PHASE1.
SOURCE malaria parasite P. falci-parum.
ORGANISM Plasmodium falci-parum
REFERENCE 1 (bases 1 to 199551)
AUTHORS Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
Hyman,R.W., Fung,E.L., Qin,F., Rowley,D., Mao,J., Tamaki,T.,
Kudr,O.B., Conway,A.B. and Davis,R.W.
TITLE Plasmodium falci-parum 3D7 chromosome 12
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2002, 22:12:43 ; Search time 137.032 Seconds
(without alignments)
7805.766 Million cell updates/sec

Title: us-09-881-556-4
Perfect score: 623
Sequence: 1 cctcaatggtgaacatcaa.....aaaaaaaaaaaaaactcgag 623

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_032802.*
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	52	8.3	6183	24	ABL33197
C 2	52	8.3	9539	22	AAS45347
C 3	51.8	8.3	6915	24	ABL332406
C 4	50.8	8.2	6494	24	ABL333393
C 5	50.2	8.1	1757	24	ABL34421
C 6	50.2	8.1	7306	24	ABL333637
C 7	49.8	8.0	14006	24	ABL33958
C 8	49.5	8.0	5493	24	AAS61386
C 9	49	7.9	5283	22	AAS46539
C 1	52	8.3	6183	24	ABL33197
C 2	52	8.3	9539	22	AAS45347
C 3	51.8	8.3	6915	24	ABL332406
C 4	50.8	8.2	6494	24	ABL333393
C 5	50.2	8.1	1757	24	ABL34421
C 6	50.2	8.1	7306	24	ABL333637
C 7	49.8	8.0	14006	24	ABL33958
C 8	49.5	8.0	5493	24	AAS61386
C 9	49	7.9	5283	22	AAS46539

C 10	48.8	7.8	6145	24	ABL32972	Human immune syste
C 11	48.6	7.8	10619	22	AAS45296	Chemically pretrea
C 12	48.6	7.8	10619	24	ABL32070	Human immune syste
C 13	48.4	7.8	3567	21	AAA70117	Plasmodium falcipa
C 14	48.4	7.8	5452	24	ABL33149	Human immune syste
C 15	48.4	7.8	6767	22	AAS46608	Tumour suppressor
C 16	48.2	7.7	6485	22	AAS46559	Tumour suppressor
C 17	48.2	7.7	6485	24	ABL33808	Human immune syste
C 18	48.2	7.7	7479	24	AAS63345	Chemically pretrea
C 19	48	7.7	7924	24	ABL34139	Human immune syste
C 20	47.8	7.7	8254	24	ABL32864	Human immune syste
C 21	47.8	7.7	13249	24	ABL32116	Human immune syste
C 22	47.6	7.6	819	22	AA194974	Human neuroblastom
C 23	47.6	7.6	8634	24	ABL33056	Human immune syste
C 24	47.2	7.6	7134	24	ABL32482	Human immune syste
C 25	47	7.5	6035	24	ABL33525	Human immune syste
C 26	47	7.5	6277	22	AAS46728	Tumour suppressor
C 27	47	7.5	6277	24	ABL34040	Human immune syste
C 28	46.8	7.5	5379	24	ABL33677	Human immune syste
C 29	46.8	7.5	5379	24	ABL34577	Human metastasis a
C 30	46.8	7.5	5413	22	AAS46694	Tumour suppressor
C 31	46.8	7.5	5454	21	AAA70236	Plasmodium falcipa
C 32	46.8	7.5	6070	24	ABL33679	Human immune syste
C 33	46.8	7.5	6070	24	ABL34579	Human metastasis a
C 34	46.8	7.5	6591	22	AAS46283	Tumour suppressor
C 35	46.8	7.5	7571	24	ABL32527	Human immune syste
C 36	46.8	7.5	9209	24	ABL34427	Human immune syste
C 37	46.8	7.5	9219	22	AAS46808	Tumour suppressor
C 38	46.8	7.5	14429	24	ABL34242	Human immune syste
C 39	46.8	7.5	17934	24	ABL33719	Human immune syste
C 40	46.6	7.5	5893	24	ABL32859	Human immune syste
C 41	46.6	7.5	6419	24	ABL32267	Human immune syste
C 42	46.4	7.4	4661	24	AAS61136	Human gene regulat
C 43	46.4	7.4	13606	22	AAS45457	Chemically pretrea
C 44	46.4	7.4	13606	22	AAS46561	Tumour suppressor
C 45	46.4	7.4	13606	24	ABL33810	Human immune syste

ALIGNMENTS

RESULT 1

ABL33197/C	ABL33197 standard; DNA; 6183 BP.
XX	AC ABL33197;
XX	DT 26-MAR-2002 (first entry)
XX	DE Human immune system associated gene SEQ ID NO: 1170.
XX	Human immune system disease; cytosine methylation; antiasthmatic;
KW	antiarteriosclerotic; antianemic; cytostatic; nootropic;
KW	neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW	antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW	antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW	acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW	neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
XX	gene; ds.
XX	OS Homo sapiens.
XX	PN WO200200928-A2.
XX	PD 03-JAN-2002.
XX	PF 02-JUL-2001; 2001WO-EP07537.
XX	PR 30-JUN-2000; 2000DE-1032529.
PR	01-SEP-2000; 2000DE-1043826.
XX	PA (EPIG-) EPIGENOMICS AG.
XX	

Db	2329	TAAGAAAAATAATATCCAAAAACAATAAAAAAATAACCAAAACCGTAATAAACAACAAA	2270
Qy	545	CATTGATAAATATGGCCCTTAGCCATTTTCCAATATCTAAATGTGCAACTCAAGTAAAAAA	604
Db	2269	AACCATTAACCAAAAACCTACCGTACCTTAAAAAAACCAATACTAATCTAATTAACCTAA	2210
Qy	605	AAAAAAAACCT 619	
Db	2209	CAAAAACATAAAAAAT 2195	
RESULT 4			
ABL33393/c			
ID	ABL33393	standard; DNA; 6494 BP.	
XX	AC	ABL33393;	
XX	AC		
DT	26-MAR-2002	(first entry)	
XX	DE	Human immune system associated gene SEQ ID NO: 1366.	
XX	KW	Human; immune system disease; cytosine methylation; antiasthmatic;	
KW	KW	antiarteriosclerotic; antianemic; cytosolic; neurotropic;	
KW	KW	neuroprotective; anti-HIV; anticonvulsant; ophthalmological;	
KW	KW	antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;	
KW	KW	acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;	
KW	KW	neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;	
KW	KW	gene; ds.	
OS	XX	Homo sapiens.	
XX	XX		
PN	WO200200928-A2.		
XX	XX		
PD	03-JAN-2002.		
XX	XX		
PF	02-JUL-2001; 2001WO-EP07537.		
XX	XX		
PR	30-JUN-2000; 2000DE-1032529.		
PR	01-SEP-2000; 2000DE-1043826.		
XX	XX		
PA	(EPIG-) EPIGENOMICS AG.		
XX	XX		
PI	Olek A, Piepenbrock C, Berlin K;		
XX	XX		
DR	WPI; 2002-130909/17.		
XX	XX		
PT	Nucleic acid comprising fragment of chemically modified gene, useful		
PT	for diagnosis and treatment of diseases associated with abnormal		
PT	cytosine methylation		
XX	XX		
PS	Claim 1; SEQ ID NO 1366; 32pp + Sequence Listing; German.		
XX	XX		
CC	The present invention provides a number of human immune system associated		
CC	genes which are modified by the methylation of cytosines. The sequences		
CC	can be used in the diagnosis and treatment of immune system disorders,		
CC	including eye diseases such as retinopathy, neovascular glaucoma and		
CC	macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid		
CC	leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,		
CC	rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel		
CC	diseases. The present sequence is a gene of the invention.		
XX	XX		
SQ	Sequence 6494 BP; 1929 A; 81 C; 1179 G; 3305 T; 0 other;		
	Query Match	8.2%; Score 50.8; DB 24; Length 6494;	
	Best Local Similarity	47.4%; Pred. No. 0.026;	
	Matches 184; Conservative	0; Mismatches 202; Indels 2; Gaps 1;	
Qy	237	TCAGATCACACTCTCTACTTCTCCAGAGAGAGGATTGACTCAGACGTTTACCATCAC	296
Db	4784	TCACAAACAACTTCAACCAATTAATAAAAAATAACAAATTTATAAAAAACAATAAT	4725
Qy	297	TTCTACAGAACAAACCTCAAAGCCGCTACATATCTTAACCAAGAAAAATGAAAAATCATAT	356

Db 4724 ATACCAATAAAACCTCTAAATAACAAAACTTAAATAAAATCTTTTATCTACAT 4665
QY 357 AATCAGGAAGTAATTCGAGTTTATACACTAAATCT--TCCAAATCAACCAATCTTT 414
Db 4664 ATTAATAATATAATTTATTTATTTCTTAAACAATTTTCTCTAAATTTCTCAAACTAT 4605
QY 415 CAGGCAGCTGGCAGATGGGAAAGTTATGGGGAAGATAAATGGTTGATGGTGAGC 474
Db 4604 CAATATATTTAAAAAAAATTTAAATTTATCTACAACTTCTCTAAATTTATATATAATAA 4545
QY 475 AAGTAATCACTAAATGACATTAATGAGATACCTCGAAGCAAAATAGAAACAAATCT 534
Db 4544 ATAAATAATACCAATATAATATAATTAACAACTTAAAAAAAACAATACATAATAAACA 4485
QY 535 AAGTATCTGCTATGATTAATATGCGCTTTAGCCATTTCCAAATATCTAAATTTGCAACTC 594
Db 4484 AATATCTTATATTCAAATAAATAAATAATAAATACATCTCTCACTAAATAATAAACA 4425
QY 595 AAGTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 622
Db 4424 CAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4397

RESULT 5
ABL34421/c
ID ABL34421 standard; DNA; 1757 BP.
XX
AC ABL34421;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 2394.
XX
KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200200928-A2.
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP07537.
XX
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-130909/17.
XX
PT Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation -
XX
PS Claim 1; SEQ ID NO 2394; 32pp + Sequence Listing; German.
XX
CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.

CC diseases. The present sequence is a gene of the invention.
SQ Sequence 1757 BP; 396 A; 83 C; 464 G; 813 T; 1 other;
Query Match 8.1%; Score 50.2; DB 24; Length 1757;
Best Local Similarity 60.7%; Pred. No. 0.027; 53; Indels 0; Gaps 0;
Matches 82; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
QY 483 ACTAAATAATGACATTTATTCGAAGATCTTCGAAAGAAATTTAGAAACAAATGTAAAGTATC 542
Db 1683 ATTAAATAATTTACATTTAAATAATAATCCAAACAAATAATAATAATAATAACTAGC 1624
QY 543 TGCATTGATAAATATGCGCTTAGCCATTTCCAAATATCTAAATTTGTCAACTCAAGTAAA 602
Db 1623 ATATCTCAAAATTAATCTAAAAAATATCTAAATAACAAATTTTACTAATAAATAAAT 1564
QY 603 AAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 617
Db 1563 ATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1549
RESULT 6
ABL33637/c
ID ABL33637 standard; DNA; 7306 BP.
XX
AC ABL33637;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 1610.
XX
KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200200928-A2.
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP07537.
XX
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-130909/17.
XX
PT Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation -
XX
PS Claim 1; SEQ ID NO 1610; 32pp + Sequence Listing; German.
XX
CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.

SQ Sequence 7306 BP; 2019 A; 93 C; 1645 G; 3549 T; 0 other;

XX WPI: 2002-017470/02.
XX
XX New nucleic acid sequences from chemically modified genes associated
PT with gene regulation, useful for analysing cytosine methylations for
PT diagnosis and therapy of diseases e.g. severe combined immunodeficiency
PT disease -
XX
XX
PS Disclosure: SEQ ID No 349; 26pp; English.
XX
XX The invention relates to 224 nucleic acid sequences comprising at least
CC 18 bases of a chemically pretreated gene associated with gene regulation
CC selected from 43 known genes (or complementary sequences). The
CC chemical pretreatment converts cytosine bases unmethylated at the
CC 5-position to uracil or another base with hybridisation behaviour
CC dissimilar to cytosine, to enable analysis of cytosine methylations.
CC The DNA sequences, oligomers (or sets/arrays) and method are
CC useful in the diagnosis of diseases (or predisposition to diseases)
CC associated with gene regulation and in therapy of such diseases, by
CC enabling analysis of the cytosine methylation patterns of such genes,
CC kits are provided. They are especially useful in diagnosis
CC and therapy of e.g. severe combined immunodeficiency disease, cardiac
CC disorders, haemophilia, solid tumours and cancer, Werner syndrome,
CC asthma, HDR syndrome, Saethre-Chotzen syndrome, renal disease,
CC precociousness, graft versus-host disease. The present sequence is a
CC sequence included in the sequence data for this specification and is
CC associated with the human gene regulation-associated genes.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX
XX Query Match 8.0%; Score 49.6; DB 24; Length 5493;
Best Local Similarity 47.4%; Pred. No. 0.048;
Matches 148; Conservative 0; Mismatches 164; Indels 0; Gaps 0;
QY 306 ACAACTCAAGCCAGTACATATCTTACCAAGAAATGAAATCATATAATCAGGAA 365
DB 4424 AAAAAACAAAAAATACAACTTAAAAAATAAAAAACAAAAAATAAAAAA 4365
QY 366 GGTAAATTCGAGTTTAAATACATCAATCTTCCAAATCAACCCAAATCTTTCAGCGAGCTG 425
DB 4364 AAAAAATATAATATACAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 4305
QY 426 GCAGATGTGGAAGTTATGGGAAAGGATATAAATGGTTGATGGTGAAGTAACTACT 485
DB 4304 AAAAAACAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 4245
QY 486 AAAAAATGACATTATGAAGTACTTCGAAAGAAATTAGAAACAAATGTAAGTATCTGC 545
DB 4244 ACAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 4185
QY 546 ATTGATAAATATGGCTTATAGCCATTTCAAAATATCTAAATTTGTCACATCAAGTAAAAAA 605
DB 4184 ATTAATAAACCAACCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 4125
QY 606 AAAAAAATAAAAA 617
DB 4124 AAAAAAATAAAAA 4113
RESULT 9
AAS46539/c
ID AAS46539 standard; DNA; 5283 BP.
XX
AC AAS46539;
XX
DT 18-DEC-2001 (first entry)
XX
DE Tumour suppressor gene derived chemically modified sequence #261.
XX

KW Human; tumour suppressor gene; oncogene; antitumour; cytostatic;
KW cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
KW cytosine methylation; ds.
OS Homo sapiens.
XX WO200168912-A2.
XX
XX 20-SEP-2001.
XX
XX 15-MAR-2001; 2001WO-EP02955.
XX
XX 15-MAR-2000; 2000DE-1013847.
PR 06-APR-2000; 2000DE-1019058.
PR 30-JUN-2000; 2000DE-1019173.
PR 01-SEP-2000; 2000DE-1032529.
XX 01-SEP-2000; 2000DE-1043826.
XX (EPIG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX WPI: 2001-602752/68.
XX
XX Fragments of chemically modified genes associated with tumour suppressor
PT genes and oncogenes, useful in designing primers and probes for
PT analysing diseases associated with cytosine methylation state e.g.
PT cancer -
XX
XX Claim 1; SEQ ID No 261; 27pp; English.
XX
XX The invention relates to a nucleic acid comprising a sequence of 18
CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
CC bisulphite, of genes associated with tumour suppression and
CC oncogenes having a sequence taken from 536 (actually 533 since
CC numbers 408, 458 and 500 are missing from the sequence listing) sequences
CC (SS) and sequences complementary to (SS). The nucleic acid may be a
CC peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may
CC form part of a set of probes for detecting the cytosine methylation state
CC and/or single nucleotide polymorphisms and also to be used in an
CC array for analysing diseases associated with CpG dinucleotides e.g.
CC cancers and tumours. The probes can also be used in a method for
CC ascertaining genetic and/or epigenetic parameters for the diagnosis
CC and/or therapy of existing diseases or the predisposition to specific
CC diseases, by analysing cytosine methylations. The parameters may be
CC compared to another set of genetic and/or epigenetic parameters, the
CC differences serving as basis for diagnosis and/or prognosis events which
CC are disadvantageous to patients. The present sequence is one of the
CC 533 genomic sequences derived from tumour suppressor genes and
CC oncogenes.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 5283 BP; 1113 A; 151 C; 1299 G; 2720 T; 0 other;
XX
XX
XX Query Match 7.9%; Score 49; DB 22; Length 5283;
Best Local Similarity 47.3%; Pred. No. 0.066;
Matches 148; Conservative 0; Mismatches 165; Indels 0; Gaps 0;
QY 305 AACAACTCAAGCCAGCTACATATCTTCAACAGAAATGAAATCATATAATCAGGA 364
DB 2358 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 2299
QY 365 AGGTAATTCGAGTTTAAATACATCAATCTTCCAAATCAACCCAAATCTTTCACCGAAGCT 424
DB 2298 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 2239
QY 425 GGCAGATGTGGAAGTTATGGGAAAGGATATAAATGGTTGATGGTGAAGTAACTAC 484
DB 2238 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 2179

QY 485 TAAAAATGACATTTATGAGTACTTCGAAGAATAATTAGAACCAAAATGTAAAGTATCTG 544
Db 2178 CATCAATCCACCACCGGATACACTCAACCTAAATAACAAAAAACTATCTCAAA 2119
QY 545 CATTTGATAATATGCGCTTACGCCATTTCCTCAATATCTTAAATTTGTCAACTCAAGTAAAAAA 604
Db 2118 AAAAAAATAATATCAATATATACTATCACTATATATCCACACTTTAAAAAAC 2059
QY 605 AAAAAAATAAATA 617
Db 2058 CAAAAAATAAATA 2046

RESULT 10
ABL32972/c
ID ABL32972 standard; DNA; 6145 BP.
XX
AC ABL32972;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 945.
XX
KW Human; Immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianemic; cytosine; cytosine; cytosine;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmologic;
KW antithrombotic; antiarthritic; antidiabetic; antipsoriasis;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200200928-A2.
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP07537.
XX
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-130905/17.
XX
PT Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation -
XX
XX Claim 1; SEQ ID NO 945; 32pp + Sequence Listing; German.
XX
CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.
XX
SQ Sequence 6145 BP; 1493 A; 99 C; 1309 G; 3244 T; 0 other;

Query Match 7.88; Score 48.8; DB 24; Length 6145;
Best Local Similarity 47.28; Pred. No. 0.076;
Matches 149; Conservative 0; Mismatches 167; Indels 0; Gaps 0;

QY 301 ACAGACAACTCAAGCCAGCTACATATCTCAACCAAGAAATGAAAAATCATATAATC 360
Db 5661 AATAATATATACAAAAAACAATAATTAATAAAAAAATAAAAAAATAAAAAAATA 5802

QY 361 AGGAAGTAAATTCGAGTTTAAATACACTAAATCTTCCAAATCAACCAATCTTTTCAGCGCA 420
Db 5801 AATAACTAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 5742
QY 421 AGCTGGCAGATGTGGAAGTTATTCGGGAAAGGATNAATGGTTGATGGTGAGCAAGTAA 480
Db 5741 AATAAAAAATAAAAAAACGATTAACCCATTAATAAAAAAATAAAAAAATAAAAAAATAAAAA 5682
QY 481 TCACATAAAATGACATTTATTCAAGATATCTTCGAAAGAAATTTAGAAACAAATGTAAAGTA 540
Db 5681 TAAATAAATAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 5622
QY 541 TCTGCATTGATAATATATGCGCTTAGCCATTTCCTCAATATCTTAAATTTGTCMACTCAAGTAA 600
Db 5621 AATAAATAAATAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 5562
QY 601 AAAAAAATAAATAAATA 616
Db 5561 AATAAATAAATAAATAAATA 5546

RESULT 11
AAS45296/c
ID AAS45296 standard; DNA; 10619 BP.
XX
AC AAS45296;
XX
DT 18-DEC-2001 (first entry)
XX
DE Chemically pretreated genomic DNA associated with cell cycle #1.
XX
KW Cell cycle; human; CpG dinucleotide; cytosine methylation; HIV; aging;
KW human immunodeficiency virus; neurodegenerative disorder; solid tumour;
KW graft-versus-host disease; glomerular disease; low body disease; cancer;
KW arthritis; arteriosclerosis; anti-HIV; neuroprotective; antiarthritic;
KW immunosuppressive; antitumour; cytostatic; antiarteriosclerotic; ds;
KW PCR primer.
XX
OS Homo sapiens.
XX
PN WO200168911-A2.
XX
PD 20-SEP-2001.
XX
PF 15-MAR-2001; 2001WO-EP02945.
XX
PR 15-MAR-2000; 2000DE-1013847.
PR 06-APR-2000; 2000DE-1019058.
PR 07-APR-2000; 2000DE-1019173.
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2001-602751/68.
XX
PT Designing primers and probes for analysing diseases associated with
PT cytosine methylation state e.g. arthritis, cancer, aging,
PT arteriosclerosis comprising fragments of chemically modified genes
PT associated with cell cycle -
XX
XX Claim 1; SEQ ID NO 1; 28pp; English.
XX
CC Sequences AAS45296-AAS45520 represent chemically pretreated genomic DNA
CC molecules associated with the cell cycle and specific PCR primers of the
CC invention. The sequences are useful for detecting the methylation state
CC of all CpG dinucleotides in a sequence and therefore for analysing
CC associated diseases. By analysing cytosine methylations in the pretreated
CC DNA, genetic and/or epigenetic parameters for the diagnosis and therapy
CC of existing diseases or the predisposition to specific diseases can be

PT Plasmodium falciparum, useful as antimalarial vaccines and in the
PT diagnosis of P.falciparum infection -
XX
PS Disclosure: Page 470-471; 577pp; English.
XX
CC The present invention describes proteins and their fragments (I) encoded
CC by chromosome 2 of the human malarial parasite, Plasmodium falciparum.
CC Also described are: (i) nucleotide sequences (II) encoding (I); and (2)
CC vaccines against P. falciparum infection comprising (I) or (II).
CC (1) and (II) are useful for the development of vaccines against
CC P. falciparum infection. (i) and polyclonal antisera or a monoclonal
CC antibody raised to immunogens comprising the sequences of (I), are
CC useful in the detection of infection with P. falciparum. Furthermore,
CC (1) especially when they are rifins or secreted or membrane proteins)
CC can aid the identification of drugs to treat or prevent P. falciparum
CC infection, or they can be used to identify drug resistance in
CC P. falciparum. Sequencing of the Plasmodium chromosome 2 and the
CC subsequent identification of proteins encoded by it will help to expand
CC our understanding of parasite biology, a process hampered by the
CC complexity of the parasitic lifecycle, and provide new targets for
CC vaccine and drug development. Parasite resistance to drugs and mosquito
CC resistance to insecticides have led to a resurgence of malaria in many
CC parts of the world, and there is a pressing need for vaccines and new
CC drugs. AAA70078 and AAB18144 to AAB18352 represent nucleotide
CC and protein sequences given in the present invention, but which are not
CC specifically mentioned within the specification.
XX
SQ Sequence 3567 BP; 1641 A; 269 C; 473 G; 1184 T; 0 other;
Query Match 7.8%; Score 48.4; DB 21; Length 3567;
Best Local Similarity 47.9%; Pred. No. 0.083;
Matches 139; Conservative 0; Mismatches 151; Indels 0; Gaps 0;
Qy 328 ATCTTAACCAAGAAATGAAATCATATATATCATGAGGTAATTCGAGTTTAAATACAC 387
Db 1373 ATATTAATTAATGATGATATATTAATTAATGATGATATTAATTAATGATGATG 1432
Qy 388 TAAATCTTCCAAATCAACCCCAATCTTTCACCCAGCTGGCAGAGTGGAAAGTTATGGG 447
Db 1433 ATATATTAATTAATTAATGATGATATTAATTAATGATGATATTAATTAATTAATA 1492
Qy 448 AAAAGGATAAAATGGTGTGAGCAAGTAATCACTTAAATATGACATTAATTCGAAGTA 507
Db 1493 ATGATATATTAATTAATTAATGATGATATTAATTAATTAATTAATTAATTAATA 1552
Qy 508 CTTCGAAGAAATAGAAACAAATGTAAAGTATCTGCATTGATAATATGCGCTTAGCC 567
Db 1553 ATAATGATAATTAATTAATTAATTAATTAATTTCTCCGACAGGTGATAAAATCATATTGAAA 1612
Qy 568 ATTTCCAAATATCTAAATTTCAACTCAAGTAAAAAATAAAAAATAAAAA 617
Db 1613 AGATCCTTTTAAAAACACATCCCAAAATGTACAGATGTTTAAAAAGAAA 1662
RESULT 14
ABL33149/G
ID ABL33149 standard; DNA: 5452 BP.
XX
AC ABL33149;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 1122.
XX
KW Human; Immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianemic; cytostatic; nootropic;
KW neuroprotective; anti-HIV; emiconvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.

OS Homo sapiens.
XX
PN WO200200928-A2.
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP07537.
XX
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
WP; 2002-130909/17.
XX
Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation -
XX
PS Claim 1; SEQ ID NO 1122; 32pp + Sequence Listing; German.
XX
CC The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.
XX
SQ Sequence 5452 BP; 1328 A; 55 C; 1065 G; 3004 T; 0 other;
Query Match 7.8%; Score 48.4; DB 24; Length 5452;
Best Local Similarity 47.9%; Pred. No. 0.091;
Matches 139; Conservative 0; Mismatches 151; Indels 0; Gaps 0;
Qy 327 TATCTTAACCAAGAAATGAAATCATATTAATCAGGAAGTAATTCGAGTTTAAATACA 386
Db 5201 TATCTCAACAAACAAACAAACAAACAAATTAATTAATTAATTAATTAATTAATTAAT 5142
Qy 387 CTAAATCTTCCAAATCAACCAATCTTTCACGCAAGCTGGCAGATGTGGAAGTTATGGG 446
Db 5141 AAAAAACAAATACACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 5082
Qy 447 GAAAGGATAAAATGGTGTGAGCAAGTAATCACTTAAATATGACATTAATTCGAAGAT 506
Db 5081 AAAAAATCATACTACTATAAAATACATAACTAAATTAATTAATTAATTAATTAATTAAT 5022
Qy 507 ACTTCGAAAGAAATAGAAACAAATGTAAAGTATCTGCATTGATAATATGCGCTTAGC 566
Db 5021 CAATAACACACTAATCAACAAACAAATTAATTAATTAATTAATTAATTAATTAATTAAT 4962
Qy 567 CATTTCCAAATATCTAAATTTCAACTCAAGTAAAAAATAAAAAATAAAAA 616
Db 4961 AATTATCCAAAAAACAAACAAACAAACAAATTAATTAATTAATTAATTAATTAATTAAT 4912
RESULT 15
AAS46608/G
ID AAS46608 standard; DNA: 6767 BP.
XX
AC AAS46608;
XX
DT 18-DEC-2001 (first entry)
XX
DE Tumour suppressor gene derived chemically modified sequence #330.
XX
KW Human; tumour suppressor gene; oncogene; antitumour; cytostatic;
KW cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
KW cytosine methylation; ds.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2002, 22:17:03 ; Search time 31.0794 Seconds
(without alignments)
4923.835 Million cell updates/sec

Title: us-09-881-556-4

Perfect score: 623
Sequence: 1 cctcaatggtgaacatcaa.....aaaaaaaaaaaaaacctcgag 623

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA.*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	46.8	7.5	7218	1	US-08-232-463-14
C 2	43.2	6.9	5852	1	US-07-867-106-2
C 3	42.8	6.9	289	4	US-09-007-005-17
C 4	42.8	6.9	289	4	US-09-244-796-17
C 5	42.4	6.8	132331	3	US-09-128-155-16
C 6	41.2	6.6	19124	2	US-08-487-826B-13
C 7	40.4	6.5	2447	2	US-09-014-969-14
C 8	39.8	6.4	2619	5	PCT-US96-10521-17
C 9	39.8	6.4	2887	5	PCT-US96-10521-14
C 10	39.4	6.3	2649	2	US-08-718-864-1
C 11	39.4	6.3	2649	2	US-09-059-964A-1
C 12	39.4	6.3	2649	2	US-08-842-341-1
C 13	37.8	6.1	1886	4	US-08-936-165A-224
C 14	37.6	6.0	2019	3	US-09-040-843-3
C 15	37.6	6.0	2019	4	US-09-621-855-3
C 16	37.6	6.0	2593	3	US-09-040-843-1
C 17	37.6	6.0	2599	4	US-09-621-855-1
C 18	37.2	6.0	658	4	US-08-998-416-595
C 19	37	5.9	566	1	US-08-211-942-10
C 20	37	5.9	2836	3	US-08-747-221B-24
C 21	37	5.9	2836	3	US-08-747-221B-26
C 22	37	5.9	2836	4	US-09-005-051-24
C 23	37	5.9	2836	4	US-09-005-051-26
C 24	36.8	5.9	240	1	US-08-628-417-6
C 25	36.6	5.9	9636	4	US-08-323-170B-1
C 26	36.6	5.9	9636	4	US-08-954-441-1
C 27	36.2	5.8	2230	4	US-08-378-313-24

28	36.2	5.8	2672	3	US-09-214-564A-5	Sequence 5, Appli
29	36.2	5.8	2815	3	US-09-214-564A-1	Sequence 1, Appli
30	36	5.8	440	4	US-08-936-165A-24	Sequence 24, Appli
31	36	5.8	1493	1	US-08-340-820-24	Sequence 24, Appli
32	36	5.8	1493	1	US-08-593-535-24	Sequence 24, Appli
33	36	5.8	2445	6	5215909-9	Patent No. 5215909
34	35.8	5.7	2223	1	US-08-257-073-4	Sequence 4, Appli
35	35.6	5.7	1119	4	US-09-178-973B-7	Sequence 7, Appli
36	35.6	5.7	1119	4	US-09-354-243B-7	Sequence 7, Appli
37	35.4	5.7	1117	4	US-09-247-373B-33	Sequence 33, Appli
38	35.4	5.7	5852	1	US-07-867-106-2	Sequence 2, Appli
39	35.2	5.7	8920	2	US-08-446-855A-1	Sequence 1, Appli
40	35.2	5.7	8920	4	US-09-150-741-1	Sequence 1, Appli
41	35	5.6	473	1	US-08-764-100-16	Sequence 16, Appli
42	35	5.6	1066	1	US-08-157-101A-4	Sequence 4, Appli
43	35	5.6	1279	4	US-09-277-716-31	Sequence 31, Appli
44	35	5.6	4970	1	US-08-764-100-14	Sequence 14, Appli
C 45	35	5.6	4970	1	US-08-764-100-20	Sequence 20, Appli

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-Fis
US-08-232-463-14

Query Match

7.5%; Score 46.8; DB 1; Length 7218;

Best Local Similarity 0.8%; Pred. No. 0.0057;
Matches 3; Conservative 215; Mismatches 142; Indels 0; Gaps 0;

QY 262 AGAAGAAGGATCGACTCAGAACGTTACCAATCCTACAGAACAACTCAAGCCAG 321
Db 1422 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1363
QY 322 CTACATATCTTACCAAGAAATGAAATCATAATCAGGAAGTAATTCGAGTTTA 381
Db 1362 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1303
QY 382 ATACATAATCTTCAATCAACCAATCTTTCAGGCAAGCTGGCAGATGTGGAAAGTT 441
Db 1302 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1243
QY 442 ATGGGAAAAGGATAAATGGTGTGAGTGGTGCAGCAATCACTAAATGACATTTATG 501
Db 1242 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1183
QY 502 AGATATCTTCAAGAAATAGAAACAAATGTAAAGTATCTGCATTTGATAATATGGCC 561
Db 1182 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1123
QY 562 TTAGCATTTCCAATATCTAAATCTCACTCAACTCAAGTAAATAAAAAAAACTCG 621
Db 1122 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1063

RESULT 2
US-07-867-106-2/c
; Sequence 2, Application US/07867106
; Patent No. 5389526
; GENERAL INFORMATION:
; APPLICANT: Slade, Martin B
; APPLICANT: Chang, Andy C M
; APPLICANT: Williams, Keith L
; TITLE OF INVENTION: Improved plasmid Vectors for Cellular
; TITLE OF INVENTION: Slime Moulds of the Genus Dictyostelium
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Woodcock Washburn Kurtz Mackiewicz & No. 5389526ris
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/867,106
; FILING DATE: 19920625
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PJ 7187
; APPLICATION NUMBER: PCT/AU90/00530
; FILING DATE: 02-NOV-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Feeney, Joanne Longo
; REGISTRATION NUMBER: 35,134
; REFERENCE/DOCKET NUMBER: RICE-0002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5852 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ANTI-SENSE: NO

FEATURE:
; NAME/KEY: CDS
; LOCATION: 2378..5038
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2378..5038
US-07-867-106-2

Query Match 6.9%; Score 43.2; DB 1; Length 5852;
Best Local Similarity 47.6%; Pred. No. 0.044; Indels 2; Gaps 1;
Matches 159; Conservative 0; Mismatches 173; Indels 2; Gaps 1;

QY 286 TTACCAATCCTTACAGAACAACTCAAGCCAGCTACATATCTTACCAAGAA--AA 343
Db 2317 TTTTACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2258
QY 344 TGAATAATCATATAATCAGGAAGGTAATTCGAGTGTAAATACATCAATCTTCCAATCA 403
Db 2257 AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2198
QY 404 ACCCAATCTTCCACCAAGCTGCAGATCTGCAAGTGTATGGGAAAGGATAAATGCT 463
Db 2197 TAAATATAATTAATAATGTCATGCCAAACTGATAAATATTTGATATATATTCCAATAT 2138
QY 464 TGATGGTGAGCAAGTAATCAGTAAATAAATGACATTTATGGAAGTACTTCGAAAGAAATAG 523
Db 2137 TATAAATAAGGTATAATAGATAGAGATAAATAAATAAATAAATAAATAAATAA 2078
QY 524 AAACAAATGTAAGTATCTGCATTCGATTAATAATATGCGCTTAGCCATTCCTCAATATCTAA 583
Db 2077 AAAAACCACCAAGTAATTAATATTAATGAGGGTTTTTTTTTTTTTTTTTTTTTTTT 2018
QY 584 ATTGCTCACTCAAGTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 617
Db 2017 TTTTTCACAAGTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1984

RESULT 3
US-09-007-005-17
; Sequence 17, Application US/09007005B
; Patent No. 6258558
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Szostak, Richard W.
; APPLICANT: Liu, Rihe
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; TITLE OF INVENTION: FUSIONS
; FILE REFERENCE: 00786/350003
; CURRENT APPLICATION NUMBER: US/09/007,005B
; CURRENT FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 60/064,491
; EARLIER FILING DATE: 1997-11-06
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 289
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translation template
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(289)
; OTHER INFORMATION: n = A,T,C or G
US-09-007-005-17

Query Match 6.9%; Score 42.8; DB 4; Length 289;
Best Local Similarity 12.9%; Pred. No. 0.021;
Matches 36; Conservative 98; Mismatches 146; Indels 0; Gaps 0;

QY 338 AGAAATGAAATCATAATAATCAGGAAGTAATTCGAGTTTATACACATAAATCTTCC 397

QY 495 ATTATTGAAGATACCTCGAAGAAATAGAACAAATGTAAGATATCTGCATTGTATAA 554
DB 2529 ATTATTTAAATCATAGGAATTAAGTATCTTTAAATTTAAGTATCTTTTTCAAAAA 2588
QY 555 TATGGCCTTAGCCATTCCAAATATCTAAATGTCAACTCAAGTAAAGTAAAAA 614
DB 2589 CATTTTAAATAGATAAATAATATCTTAAATTTAAGTATCTTTTTCAAAAA 2648
QY 615 A 615
DB 2649 A 2649

RESULT 11
US-09-059-964A-1
; Sequence 1, Application US/09059964A
; Patent No. 5952228
; GENERAL INFORMATION:
; APPLICANT: van der Bruggen, Pierre; Mandruzzato, Susanna;
; APPLICANT: Boon-Falleur, Thierry
; TITLE OF INVENTION: Nucleic Acid Molecule Encoding A Functional
; TITLE OF INVENTION: Protein, The Protein So Encoded And Uses Thereof
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski, L.L.P..
; STREET: 555 Fifth Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/059,964A
; FILING DATE:
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/718,964
; FILING DATE: 26-September-1996
; APPLICATION NUMBER: 08/669,590
; FILING DATE: 24-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5952228man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5442.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2649 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-09-059-964A-1

Query Match 6.3%; Score 39.4; DB 2; Length 2649;
Best Local Similarity 57.9%; Pred. No. 0.31;
Matches 70; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 495 ATTATTGAAGATACCTCGAAGAAATAGAACAAATGTAAGATATCTGCATTGTATAA 554
DB 2529 ATTATTTAAATCATAGGAATTAAGTATCTTTAAATTTAAGTATCTTTTTCAAAAA 2588
QY 555 TATGGCCTTAGCCATTCCAAATATCTAAATGTCAACTCAAGTAAAGTAAAAA 614
DB 2589 CATTTTAAATAGATAAATAATATCTTAAATTTAAGTATCTTTTTCAAAAA 2648
QY 615 A 615

RESULT 12
US-08-842-341-1
; Sequence 1, Application US/08842341
; Patent No. 5955313
; GENERAL INFORMATION:
; APPLICANT: van der Bruggen, Pierre; Mandruzzato, Susanna;
; APPLICANT: Boon-Falleur, Thierry
; TITLE OF INVENTION: Nucleic Acid Molecule Encoding A Functional
; TITLE OF INVENTION: Protein, The Protein So Encoded And Uses Thereof
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/842,341
; FILING DATE: 24-APR-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/718,964
; FILING DATE: 26-September-1996
; APPLICATION NUMBER: 08/669,590
; FILING DATE: 24-JUNE-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5955313man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5442.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2649 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-842-341-1

Query Match 6.3%; Score 39.4; DB 2; Length 2649;
Best Local Similarity 57.9%; Pred. No. 0.31;
Matches 70; Conservative 0; Mismatches 51; Indels 0; Gaps 0;

QY 495 ATTATTGAAGATACCTCGAAGAAATAGAACAAATGTAAGATATCTGCATTGTATAA 554
DB 2529 ATTATTTAAATCATAGGAATTAAGTATCTTTAAATTTAAGTATCTTTTTCAAAAA 2588
QY 555 TATGGCCTTAGCCATTCCAAATATCTAAATGTCAACTCAAGTAAAGTAAAAA 614
DB 2589 CATTTTAAATAGATAAATAATATCTTAAATTTAAGTATCTTTTTCAAAAA 2648
QY 615 A 615
DB 2649 A 2649

RESULT 13
US-08-936-165A-224/c
; Sequence 224, Application US/08936165A
; Patent No. 6348582
; GENERAL INFORMATION:
; APPLICANT: Black, Michael
; APPLICANT: Burnham, Martin
; APPLICANT: Hodgson, John

```

; APPLICANT: Knowles, David
; APPLICANT: Lonetto, Michael
; APPLICANT: Nicholas, Richard
; APPLICANT: Pratt, Julie
; APPLICANT: Reichard, Richard
; APPLICANT: Rosenberg, Martin
; APPLICANT: Ward, Judith
; TITLE OF INVENTION: No. 6348582e1 Prokaryotic Polynucleotides,
; TITLE OF INVENTION: Polypeptides and Their Uses
; NUMBER OF SEQUENCES: 534
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/936,165A
; FILING DATE: 24-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/027,032
; FILING DATE: 24-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P50349
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 224:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1886 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; US-08-936-165A-224

Query Match 6.1%; Score 37.8; DB 4; Length 1886;
Best Local Similarity 44.8%; Pred. No. 0.72;
Matches 138; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

Qy 314 AAGCCAGCTACATATCTTACCAAGAAATGAAATCATATATATCATGAGGTAATTC 373
Db 1234 ACAACACCTAATGCACTGACGAGAAAGCAGGCTGCTTTAATCAATCACTCAACT 1175

Qy 374 GAGTTTAAATACATAATCTTCCAAATCAACCCCAATCTTTCAGCGAAGCTGGCAGATGT 433
Db 1174 TAAAGATCAAGCAATTAATCAAAATTAATCAAAACCAACAAATGATCAGGTAGACACAA 1115

Qy 434 GGAAGGTTATGGGAAAGGATAAATGTTGATGGTGAGCAAGTAATCACTAAATAATGA 493
Db 1114 TACAAATCAACGGTAAATGCTATAGATAATGTTGNAGCTGAAGTAGTAATTAACCCAAA 1055

Qy 494 CATTTATTGAAGATACCTTCGAAGAAATTTAGAAACAAAATGTAAAGTATCTGCATTGTATA 553
Db 1054 GGCAATTCGACATNTTGAAGAAAGCTGTTAAAGAAAGCAACAGCAAAATGATATAGTCT 995

Qy 554 ATATGGCCTTAGCCATTTCGAAATATCTAAATTTGTCAACTCAAGTAAAAAAGAAAAA 613
Db 994 TGATTCACAGATNATGAGAAAGAGTTGCTTTCACAGCATTAGCTAAAGAAAAAGAAA 935

Qy 614 AAAACTCG 621
Db 934 AGCACTTG 927
```

```

RESULT 14
US-09-040-843-3
; Sequence 3, Application US/09040843
; Patent No. 6124119
; GENERAL INFORMATION:
; APPLICANT: Jaworski, Deborah J.
; APPLICANT: Wang, Min
; APPLICANT: Shilling, Lisa K.
; APPLICANT: Burnham, Martin
; APPLICANT: Fosberry, Andrew
; APPLICANT: Hodgson, John E.
; APPLICANT: Lawlor, Elizabeth
; APPLICANT: Rosenberg, Martin
; APPLICANT: Ward, Judith
; TITLE OF INVENTION: MecB
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert, Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2793
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/040,843
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/057,535
; FILING DATE: 29-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Falk, Stephen T
; REGISTRATION NUMBER: 36,795
; REFERENCE/DOCKET NUMBER: GM10082
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-994-2488
; TELEFAX: 215-994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2019 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-09-040-843-3

Query Match 6.0%; Score 37.6; DB 3; Length 2019;
Best Local Similarity 53.4%; Pred. No. 0.83;
Matches 79; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

Qy 475 AAGTAATCACTAAAATGACATTTATGAAGATACCTTCGAAAGAAATTTAGAAACAAATGT 534
Db 764 AAGTAAGCTTTAAGAGTCATACGACACCTAATAATTTAAAGAAATTTGAACAAGAAATTTG 823

Qy 535 AAGTATCTGCAATTTGATAAATATATGCGCTTAGCCATTTAGCCAAATATCTAAATTTGTCAACTC 594
Db 824 AAAAGTAAAAATGAAAAAGATGCCGAGTACATGCTCAAGAGTTTGAAAAATGCTGCTGCTA 883

Qy 595 AAGTAAAAAAGAAAAAAGAAAAAAGAAAAAAGTCA 622
Db 884 ACCTGCGTGTATAAACAAACAAACTTGA 911

RESULT 15
US-09-621-855-3
; Sequence 3, Application US/09621855
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2002, 23:27:34 ; Search time 1099.55 Seconds
(without alignments)
7647.311 Million cell updates/sec

Title: US-09-881-556-4

Perfect score: 623

Sequence: 1 cctcaaatggtgaacatcaa.....aaaaaaaaaaactcgag 623

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_htc:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_htc:*
- 12: gb_gss:*
- 13: em_gss_hum:*
- 14: em_gss_inv:*
- 15: em_gss_pln:*
- 16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	361.8	58.1	837	12	AQ411949
2	60.8	9.8	1101	12	CNS003B0
3	60.4	9.7	813	12	CNS03CIW
C 4	59.4	9.5	897	12	CNS07ABZ
5	59.2	9.5	1007	12	CNS06X9S
C 6	58	9.3	725	12	BH180166
C 7	57.2	9.2	891	12	CNS009JU
C 8	57.2	9.2	1101	12	CNS00LTJ
9	56.8	9.1	1001	12	CNS0155H
C 10	56.2	9.0	784	9	AL525973
11	55.8	9.0	1013	12	CNS02PVS
12	55.6	8.9	576	12	CNS007G4
C 13	55.6	8.9	1101	12	CNS017V2
14	55.4	8.9	430	12	CNS008K5
15	55.2	8.9	581	12	CNS010PZ
16	55.2	8.9	781	12	CNS009DO
17	55.2	8.9	996	12	CNS009FUH

C 18	55	8.8	931	12	CNS03XCF
C 19	55	8.8	982	12	AQ325799
C 20	55	8.8	1101	12	CNS0100X
C 21	55	8.8	1135	12	CNS033GQ
C 22	54.6	8.8	959	12	CNS00655
C 23	54.2	8.7	395	12	CNS021CB
C 24	54.2	8.7	879	12	CNS01JRG
C 25	54	8.7	828	12	CNS01ITX
C 26	53.8	8.6	802	12	AZ197934
C 27	53.8	8.6	825	3	BI644509
C 28	53.8	8.6	942	10	BM416122
C 29	53.8	8.6	991	12	CNS00JW8
C 30	53.8	8.6	1101	12	CNS00Z15
C 31	53.6	8.6	796	12	CNS0118D
C 32	53.4	8.6	614	12	CNS0152H
C 33	53.4	8.6	737	9	AI525942
C 34	53.4	8.6	807	12	CNS040I3
C 35	53.4	8.6	1043	12	CNS0145P
C 36	53.4	8.6	1203	12	CNS015WU
C 37	53.4	8.6	1380	10	BG563591
C 38	53.2	8.5	881	10	BF265362
C 39	53	8.5	900	12	CNS01574
C 40	53	8.5	1101	12	CNS01844
41	52.8	8.5	855	12	AZ183849
42	52.6	8.4	811	12	CNS007W9
C 43	52.6	8.4	867	12	CNS0054A
C 44	52.6	8.4	870	12	AQ866797
C 45	52.6	8.4	871	12	AZ671726

ALIGNMENTS

RESULT 1

AQ411949

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

AQ411949 837 bp DNA linear GSS 12-MAY-2000
CpG0325A CpIOWAGDNA1 Cryptosporidium parvum genomic, DNA sequence.

AQ411949
CpG0325A CpIOWAGDNA1 Cryptosporidium parvum genomic, DNA sequence.

AQ411949.1 GI:4469673
GSS.

Cryptosporidium parvum.
Cryptosporidium parvum.

Eukaryota; Alveolata; Apicomplexa; Coccidia; Elmeritida;
Cryptosporididae; Cryptosporidium.

Strong, W.B. and Nelson, R.G.
Preliminary profile of the Cryptosporidium parvum genome: an

expressed sequence tag and genome survey sequence analysis
Mol. Biochem. Parasitol. 107 (1), 1-32 (2000)

20183851
Contact: Nelson, R. G.
Depts. of Medicine & Pharmaceutical Chemistry

San Francisco General Hospital-University of California, San
Francisco

Box 0811, San Francisco, CA 94143-0811, USA
Tel: 415 206 8846

Fax: 415 206 3353
Email: malariad@itsa.ucsf.edu

For Annotation Data see http://medasfgh.ucsf.edu/ld/CpTags/home.html
Seq primer: M13(-21) forward

Class: shotgun.
Location/Qualifiers

1. 837
/organism="Cryptosporidium parvum"

/strain="IOWA"

/db_xref="taxon:5807"

/clone_lib="CpIOWAGDNA1"

/lab_host="E. coli XL2 Blue MRF"

/note="Vector: pBlueScript II (SK-); Site 1: EcoRV; C.
parvum (IOWA isolate) genomic DNA was hydrodynamically

sheared to produce fragments having a tight size
distribution between 2-4 kb by Dr. Irvonne Thorstenson of
the Stanford DNA Sequencing and Technology Center

(http://sequence-www.stanford.edu/group/techdev/shear.htm). The randomly sheared gDNA was chromatographed on Sephacryl S-400 to remove any small fragments and DNA eluting in the void volume was subcloned into an EcoR V-digested, alkaline phosphatase-treated pBluescript II (SK-) vector and transformed into E. coli strain XL2 Blue MRF'. Recombinant clones from the first plating of the library were selected for sequence analysis using T3 and T7 primers."

BASE COUNT 333 a 128 c 124 g 252 t
ORIGIN
Query Match 58.1%; Score 361.8; DB 12; Length 837;
Best Local Similarity 98.8%; Pred. No. 4.5e-49;
Matches 396; Conservative 0; Mismatches 2; Indels 3; Gaps 3;

QY 205 CTCTGCTACTATACAGAAACAGGATCCAGCTCAGATCACACT-CTGCTACTTCTCCAG 263
|||||
Db 13 CTCTGCTACTATACAGAAACAGGATCCAGCTCAGATCACACTCTGCTACTTCTCCAG 72
QY 264 AAGAAGGATTGGACTCAGAACG-TTACCAATCACTTCTACGAAC-AAACTCAAAAGCCAG 321
|||||
Db 73 AAGAAGGATTGGACTCAGAACGCTTACCAATCACTTCTACGAACAAAACTCAAAAGCCAG 132
QY 322 CTACATATCTTACCCAGAAAATGAAATCATTAATCAGGAAGGTAATTCGAGTTTGA 381
|||||
Db 133 CTACATATCTTACCCAGAAAATGAAATCATTAATCAGGAAGGTAATTCGAGTTTGA 192
QY 382 ATACACTTAATCTTCCAAATCAACCAATCTTTCAGCAAGCTGCGAGATGTGGAAGTT 441
|||||
Db 193 ATACACTTAATCTTCCAAATCAACCAATCTTTCAGCAAGCTGCGAGATGTGGAAGTT 252
QY 442 ATGGGAAAAGGATAAAATGGTTGATGGTGAGCAAGTAATCACTAAAATGACATTATTG 501
|||||
Db 253 ATGGGAAAAGGATAAAATGGTTGATGGTGAGCAAGTAATCACTAAAATGACATTATTG 312
QY 502 AGATACCTCGAAGAAATTAAGAACAAATGTAAGTATCTGCATTGTAATATATGCCC 561
|||||
Db 313 AGATACCTCGAAGAAATTAAGAACAAATGTAAGTATCTGCATTGTAATATATGCCC 372
QY 562 TTAGCCATTTCCTCAATATCTAAATTTGTCAACTCAAGTAAAA 602
|||||
Db 373 TTAGCCATTTCCTCAATATCTAAATTTGTCAACTCAAGTAAAA 413
|||||

RESULT 2
CNS003B0 1101 bp DNA linear GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
DEFINITION BACR08E08 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION AL064078.1 GI:4941834
VERSION
KEYWORDS GSS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 1101)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mamoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,

NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.bufile.edu/drosophila_bac.htm.

FEATURES
Location/Qualifiers
1..1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="RPCI-98"
/clone="BACR08E08"
/note="end : TET3"
BASE COUNT 553 a 143 c 100 g 134 t 171 others
ORIGIN

Query Match 9.8%; Score 60.8; DB 12; Length 1101;
Best Local Similarity 38.3%; Pred. No. 0.75;
Matches 159; Conservative 42; Mismatches 214; Indels 0; Gaps 0;
QY 203 ATCTTCTGCTACTATACAGAAACAGGATCCAGCTCAGATCACACTCTGCTACTTCTCCA 262
|||||
Db 451 ANKKKKKBYTTTTTCCNNAAYBYBYTCCCAAAAGAAAGCAAAACCCMA 510
QY 263 GAAGAAGGATTGGACTCAGAACGTTTACCAATCACTTCTACAGCAAACTCAAAAGCCAG 322
|||||
Db 511 AAAAAAAMAMAMAMMCMCCCAACCAAAATACAAAMAAACCCCAAAACMAA 570
QY 323 TACATATCTTACCAAGAAATGAAATCATTAATCAGGAAGTAATTCGAGTTTAA 382
|||||
Db 571 AAAAAAMMCCCCCMMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAA 630
QY 383 TACACTAATCTTCCAAATCAACCAATCTTTCAGCAAGCTGCGAGATGTGGAAGTTA 442
|||||
Db 631 CAAAAAATTTAAAAAAMAMMCMCCCAACCCCMCMCCCMCMCMCMCMCMCMCMCMCMCM 690
QY 443 TGGGAAAAGGATAAAATGGTTGATGGTGAGCAAGTAATCACTAAAATGACATTATTGA 502
|||||
Db 691 WAAANNAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAA 750
QY 503 AGATACCTCGAAGAAATTAAGAACAAATGTAAGTATCTGCATTGTAATATGCGCT 562
|||||
Db 751 AAAAAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAA 810
QY 563 TAGCCATTTCCTCAATATCTAAATTTGTCAACTCAAGTAAAAAAMMAAAMMAAAMMAA 617
|||||
Db 811 WAAATTTWAAAAAAMWRTTTTTTAAWATATNAAAAAAMMAAAMMAAAMMAAAMMAA 865
|||||

RESULT 3
CNS03CIW 813 bp DNA linear GSS 15-MAY-2000
LOCUS Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
DEFINITION 04B14 of library G from Tetraodon nigroviridis, genomic survey sequence.

ACCESSION AL237857
VERSION AL237857.1 GI:7896992
KEYWORDS GSS: genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorphi; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.

REFERENCE 1 (bases 1 to 813)
AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.
TITLE Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 813)

AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Bounneau,L., Fisher,C., Bernot,A., Fzimes,C., Wincker,P., Brottier,P., Quetier,F., Saurin,W. and Weissbach,J.

TITLE Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence

JOURNAL Unpublished

REFERENCE 3 (bases 1 to 813)

AUTHORS Genoscope.

TITLE Direct Submission

JOURNAL Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases

COMMENT This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetraodon>.

FEATURES

source

1. .813

/organism="Tetraodon nigroviridis"

/db_xref="taxon:99883"

/clone_lib="G"

/note="Genoscope sequence ID : C0BG014DA07SP1-end : PUC-ori"

BASE COUNT 350 a 152 c 149 g 119 t 43 others

ORIGIN

Query Match 9.7%; Score 60.4; DB 12; Length 813;

Best Local Similarity 44.5%; Pred. No. 0.97;

Matches 137; Conservative 22; Mismatches 149; Indels 0; Gaps 0;

QY 309 AACTCAAGCCAGCTACATCTCTCAACCAAGAAATCAAAATCATATATATCAGGAGGT 368

Db 128 MHHMMVMMVMMCMCTAAAGATATAAAAAAATAAAAAAATAAAAAAATAAAAAA 187

QY 369 AATTCGAGTTTAAATACACTAAATCTTCCAAATCAACCCCAATCTTTCACGCAAGTGGCA 428

Db 188 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 247

QY 429 GATGTGGAAGTATGGGGAAGGATAAAATGGTTGATGGTGAGCAAGTAATCACTAAA 488

Db 248 TAAGAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 307

QY 489 AATGACATATTGAGATCTCTCGAAGAAATAGAACCAAAATGTAAAGTATCTGCATT 548

Db 308 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAA 367

QY 549 GATAAATATGGCCTTAGCCATTTCCAAATATCTAAATTTGTCAACTCAAGTAATAAAAAA 608

Db 368 CATAAATATAATCAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAA 427

QY 609 AAAAAAA 616

Db 428 AAAAAAA 435

RESULT 4

CNS07ABZ/c 897 bp DNA linear GSS 08-JUL-2001

LOCUS T7 end of clone BC0AA003A12 of library BC0AA from strain CBS 767 of

DEFINITION Debaromyces hansenii, genomic survey sequence.

ACCESSION AL436389

VERSION Debaromyces hansenii

KEYWORDS GSS.

SOURCE AL436389.1 GI:12219802

ORGANISM Debaromyces hansenii.

Debaromyces hansenii

Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

Saccharomycetales; Saccharomycetaceae; Debaromyces.

1 (bases 1 to 897)

Lepingle,A., Casaregola,S., Neuvéglise,C., Bon,E., Nguyen,H., Artiguenave,F., Wincker,P. and Gaillardin,C.

TITLE Genomic exploration of the hemiascomycetous yeasts: 14.

JOURNAL Debaromyces hansenii var. hansenii

MEDLINE FEBS Lett. 487 (1), 82-86 (2000)

20584724

2 (bases 1 to 897)

Souciot,J.L., Aigle,M., Artiguenave,F., Blandin,G., Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S., de-Montigny,J., Dujon,B., Durres,P., Lepingle,A., Llorente,B., Malpertuy,A., Neuvéglise,C., Ozier-Kalogeropoulos,O., Potier,S., Saurin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M., Wincker,P. and Weissbach,J.

TITLE Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies

JOURNAL FEBS Lett. 487 (1), 3-12 (2000)

20584711

REFERENCE 3 (bases 1 to 897)

AUTHORS Genoscope.

TITLE Direct Submission

JOURNAL Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage, 2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

COMMENT This GSS is part of a random genomic sequencing program of thirteen yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii, Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia angusta, Debaromyces hansenii var. hansenii, Pichia sorbitophila, Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.

FEATURES

Location/Qualifiers

1. .897

/organism="Debaromyces hansenii"

/strain="CBS 767"

/variety="hansenii"

/db_xref="taxon:4959"

/clone_lib="BC0AA"

/note="end : T7"

BASE COUNT 148 a 70 c 47 g 590 t 42 others

ORIGIN

Query Match 9.5%; Score 59.4; DB 12; Length 897;

Best Local Similarity 45.9%; Pred. No. 1.4;

Matches 150; Conservative 13; Mismatches 164; Indels 0; Gaps 0;

QY 291 AATCAGCTTACAGAACAACTCAAGCCAGCTACATATCTCAACCAAGAAATGAAAT 350

Db 697 AATCTATTTTATATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 638

QY 351 CATAAATATCAGGAGGTAAATTCGAGTTTAAATACACTAAATCTTCCAATCAACCAAT 410

Db 637 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 578

QY 411 CTTTCAGCAAGCTGGCAGATGTGGAAGTTATGGGGAAGGATAAATGGTTGATGGT 470

Db 577 AAAAAAATAAAAAATWAAWAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 518

QY 471 GAGCAAGTAACTAAATATGACATTTATGAGATACATCTCGAAGAAATTTAGAACAAA 530

Db 517 WAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 458

QY 531 ATGTAAGTATCTGCATTGTGATAATATATGGCTTTAGCCATTTCCCAATATCTAAATTC 590

Db 457 AATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 398

QY 591 ACTCAAGTAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATA 617

Db 397 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 371

RESULT 5

CNS06X9S 1007 bp DNA linear GSS 06-JUL-2001

LOCUS T3 end of clone AX0AA039F08 of library AX0AA from strain CBS 7064

DEFINITION of Pichia farinosa, genomic survey sequence.

AL419462	AL419462.1	GI:12202640	
VERSION	GSS.		
KEYWORDS	Pichia farinosa		
SOURCE	Pichia farinosa		
ORGANISM	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Pichia.		
REFERENCE	de Montigny, J., Spehner, C., Souciet, J., Tekaia, F., Dujon, B., Wincker, P., Artiguenave, F. and Potier, S.		
AUTHORS	Genomic exploration of the hemiascomycetous yeasts: 15. Pichia sorbitophila		
TITLE	PEBS Lett. 487 (1), 87-90 (2000)		
JOURNAL	20584725		
MEDLINE	2 (bases 1 to 1007)		
REFERENCE	Souciet, J.L., Aigle, M., Artiguenave, F., Blandin, G., Bolotin-Fukuhara, M., Bon, E., Brottier, P., Casaregola, S., de-Montigny, J., Dujon, B., Durrens, P., Lepingle, A., Llorente, B., Malpertuy, A., Neuveglise, C., Ozier-Kalogenopoulos, O., Potier, S., Saurin, W., Tekaia, F., Toffano-Nioche, C., Weslowski-Louvel, M., Wincker, P. and Weissenbach, J.		
AUTHORS	Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies		
TITLE	PEBS Lett. 487 (1), 3-12 (2000)		
JOURNAL	20584711		
MEDLINE	3 (bases 1 to 1007)		
REFERENCE	Genoscope.		
AUTHORS	Direct Submission		
TITLE	Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage, 2 rue Gaston Cremlieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail: seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)		
JOURNAL	This GSS is part of a random genomic sequencing program of thirteen yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii, Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila, Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.		
FEATURES	Location/Qualifiers		
source	1. .1007		
	/organism="Pichia farinosa"		
	/strain="CBS 7064"		
	/db_xref="taxon:4920"		
	/clone="AX0AA039F08"		
	/clone_lib="AX0AA"		
	/note="end : T3"		
BASE COUNT	533 a 86 c 126 g 161 t 101 others		
ORIGIN			
Query Match	9.5%; Score 59.2; DB 12; Length 1007;		
Best Local Similarity	40.7%; Pred. No. 1.4;		
Matches 123; Conservative	45; Mismatches 133; Indels 1; Gaps 1;		
QY	314 AAAGCAGCTTACACTCCTTAACCAAGAAATGAAATCATATTAATCAGCAAGGTAATTC 373		
Db	143 AAAAAAATWAAANWTTAAWAAAAAATATAAAGTTTAAAAAGAAAGAAATGA 202		
QY	374 GAGTTTTAATACACTAAATCTTCCAATCAACCAACTCTTTTCACGCAAGCTGCAGATGT 433		
Db	203 WAAATWAAAAAAMWMAAAACAMGTRCATGTGTCWGGGTGATGAWGTATWATWANGAAW 262		
QY	434 GGAAGTTATGGGAAAAGGATAAATGTTGTATGGTGAGCAAGTATATCACTAAAAATGA 493		
Db	263 ARAWAFTAAWGGTTAAAGWMAWAAAAAAGAAAGWGGTTTATAAWAAWAAWAAWAAW 322		
QY	494 CATATTGTAAGATACCTTCGAAAGAAATATAGAAACAAATGTAAGATCTGTCATTGTAA 553		
Db	323 WAAWAAAAAAMAAACGMAAAWAAWAAAAAAMAAWAAWAAWAAWAAWAAWAAWAAW 391		
QY	554 ATATGCCTTAGCCATTTCCTCAATATCTAAATTTGTCAACTCAAGTAAAAAATAAAAA 613		

QY	332	TAACCAAGAAAAATGAAATATCATATATATCAGGAGGTAAATTCGAGTTTAAATACACTAAA	391			
Db	672	TWATAATAAATAAATBMAAAAAAAAAAAAAAAAAAAAAAAAAAAGSTSBYSBTNSAVA	613			
QY	392	TCITCCAAATCAACCCCAATCTTTCACGCAAGCTGGCAGATGTGGMAAGTTATGGGGAAA	451			
Db	612	TSWAABBAAAAAAAAAAAYSBTATBBSBBAAAAAAAAAGAGKKGACGAKRAGAAGAA	553			
QY	452	GGATAAATGGTGTGAGCAAGTAATCACTATMAAATGACATATTTGGAAGATACCTC	511			
Db	552	AGAAGAGGGRAMAMCTYACMKCTCRYTAMAAATTAACACTTAATATATAAATAATTA	493			
QY	512	GAAAGAAATTTAGAACAAAAATGTAAAGTATCTGCATGTGATAATATCGCCTTACCCATTT	571			
Db	492	AWATTAAAAAATAAATAAATCTTTHAATCYWAAAAAATAAATCAAAATMTTAAAYAYAY	433			
QY	572	CCAAATATCAATTTGTCACATCAAGTAAAAAATAAATAAATAAATAAATAAATAA	617			
Db	432	TYTAHAATAATYMAAAAAAATTAATAAATAAATAAATAAATAAATAAATAAATAA	387			
RESULT 8						
CNS00LT2/C						
LOCUS	CNS00LT2 1101 bp DNA linear GSS 14-JUN-1999					
DEFINITION	Drosophila melanogaster genome survey sequence TET3 end of BAC: BACR48P19 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.					
ACCESSION	AL078714.1 GI:5102004					
VERSION	GSS.					
KEYWORDS	fruit fly.					
SOURCE	Drosophila melanogaster					
ORGANISM	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.					
REFERENCE	1 (bases 1 to 1101)					
AUTHORS	Genoscope.					
TITLE	Direct Submission					
JOURNAL	Submitted (11-JUN-1999), Genoscope - Centre National de Sequençage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)					
COMMENT	Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammosier in Pletier de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm .					
FEATURES						
source						
1..1101						
/organism="Drosophila melanogaster"						
/db_xref="taxon:7227"						
/clone_lib="RPCI-98"						
/clone="BACR48P19"						
/note="end : TET3"						
BASE COUNT	469 a	6 c	69 g 151 t 406 others			
ORIGIN						
Query Match 9.2%; Score 57.2; DB 12; Length 1101;						
Best Local Similarity 19.9%; Prod. No. 2.8;						
Matches 116; Conservative 193; Mismatches 273; Indels 0; Gaps 0;						
QY	36	CAAGCCCTATTACCTCACCAGAAACGAATCAAGTTCAATCATCCCTCTGTGACAGTT	95			
Db	1072	MMMMMAAAMMCMCCMCCMCMCAAMMAMMTTMMMMMAAMMMMAAMMMMAAMMM	1013			

```
QY 96 CCAGTACTGGATCAGTCTCAAAATCTCCTCTGTTACTATTCCAGAGACTGGATCAGACTC 155
Db 1012 MMATTTTTHMMAMAMMMMMMMMMNATTHAHTTTTTHTTTMMAMCMCTTTHMMMMMM 953
QY 156 AGATCAGCGCCTTGAGCAATTCAGAGACTGGATCAGTCTCAAAATCATTCTTGCTACT 215
Db 952 MMAMMMMMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 893
QY 216 ATACCAAGAAACAGGATCCAGCTCAGATCAGTCTCTCTCTCTCCAGAGAAGGATGG 275
Db 892 CMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 833
QY 276 ACTCAGAAGCTTACCATCAGTCTTACAGAGCAAACTCAAGCCAGCTACATCCTAC 335
Db 832 HMMAMMYHTTMMMMTMMMMCMCMCMMAAAATMMMMMMAMHMTTHHCTMMCMCMCMCC 773
QY 336 CAAGAAATGAATCAATATATCAGGAAGTAAATTCGAGTTTAAATACATCAATCTT 395
Db 772 MCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 713
QY 396 CCAATCAACCCCAATCTTTACGCAAGCTGGCAGATGTGGAAGTATTCGGGAAAAGGAT 455
Db 712 MHTTTTMMMMMAAAWTTTTHAHAATTTTTTTTATTAATAAAAAAAAMAAWTAATAAAA 653
QY 456 AAAATGGTGTGAGGCAAGTAATCACTAAATGACATTAATGAAGATACCTCGAAA 515
Db 652 AAAATWTAWAAAAAATAATTAATAAAWAAAAAATAHWTMMMTWAMMTAAWANA 593
QY 516 GAAATTAGAACAATAATGTAAGTATCTGATTAATATGCGCTTAGCCATTTCCAA 575
Db 592 AAATWATAAAAAWTTTWTATTTATMMMAHAMATMMMAWAAAAAATAATTAATWAA 533
QY 576 ATATCTAAATCTCACTCAAGTAAAAAATAAAAAAATAAAAAA 617
Db 532 WAAWTAAAAAHWTWTMMHTTHAAAAWAAAAAATAATMMWAAAAAATA 491

RESULT 9
CNS0155H 1001 bp DNA linear GSS 26-JUL-1999
LOCUS Drosophila melanogaster genome survey sequence SP6 end of BAC
DEFINITION BACN13C23 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL105023
VERSION AL105023.1 GI:5617037
KEYWORDS fruit fly.
SOURCE Drosophila melanogaster
ORGANISM Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1001)
Genoscope.
Direct Submision
Submitted (23-JUL-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - France (E-mail : sedref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
http://www.edgp.ebi.ac.uk -. This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.
FEATURES
Location/Qualifiers
source
1..1001
/organism="Drosophila melanogaster"
/plasmid="pBelobAC11"
/db_xref="taxon:7227"
/clone_lib="DrosBAC"
/clone="BACN13C23"
```

```
BASE COUNT 266 a 219 c 134 g 150 t 232 others
ORIGIN
Query Match 9.1%; Score 56.8; DB 12; Length 1001;
Best Local Similarity 34.7%; Pred. No. 3.4; Indels 0; Gaps 0;
Matches 109; Conservative 60; Mismatches 145;
QY 304 GAACAACTCAAGCAGCTACATCTTAACCAACAAATGAAATCATATAATCAGG 363
Db 671 KAARERAAAAAARAGDAARAARAAARDAARAAAAATAAAWAAATAARWAAR 730
QY 364 AGGTAAATTCGAGTTTAAATACACTAAATCTTCCAAATCAACCCATCTTCCAGCAAG 423
Db 731 DARAAAAAATTAWTTAAWAAWAAWAAWAAATATAAATTTTAAATAAAAAAWWT 790
QY 424 TGGCAGATCTGGAAGTATATGGGAAAGGATAAAATGTTGATGTCGACGAGTAATCA 483
Db 791 AAAAAAATATTTAAWAAWAAAAATWAAAAATWAAAAAATWAAAAAATTTAAATTT 850
QY 484 CTAAAAATGACATTAATGAAGATACCTCGAAGAAATTAAGAACAAATGTAAGTATCT 543
Db 851 TTTAAWAAWAAWAAWAAWAAATATAAADAARAAAAAATAAAATWAAAAATTTW 910
QY 544 GCATGATAAATATGCGCTTACGCAATTTCCAAATATCTAAATTTGTCACACTCAAGTAAAA 603
Db 911 WTAWAWADTTTWTWTWAAWAAATAATWTAATTTAAATWTAATAWAAATAAAWATAAAAA 970
QY 604 AAAAAAATAAAAAA 617
Db 971 ATWAAWAAWAAWATA 984

RESULT 10
AL525973/C 784 bp mRNA linear EST 13-FEB-2001
LOCUS AL525973 LTI_NFL003_NBC3 Homo sapiens cDNA clone CS0DC015YJ02 5
DEFINITION prime, mRNA sequence.
ACCESSION AL525973
VERSION AL525973.1 GI:12789466
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 784)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: sedref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
Location/Qualifiers
source
1..784
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DC015YJ02"
/clone_lib="LTI_NFL003_NBC3"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/Note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
```

BASE COUNT 119 a 80 c 6 g 443 t 136 others

Matches 147; Conservative 34; Mismatches 163; Indels 4; Gaps 1;

Qy	271	ATTGACTCAGAAAGTTTACCAATCACCTTCTACAGAACAAACTCAAAGCCAGCTACATATC	330
Db	189	AWTAGGGTAAGGAAAAA : : : :	248
Qy	331	CTAACCAAGAANAATGAANAATCATTAATATCAGGAAGGTAAATTCGAGTTTTTAATACACTAA	390
Db	249	RCAWCGCGCCAAAAMAAAACAAAAAATAATATAAAAAAAAAAAAAAAAAAAAAWA	308
Qy	391	ATCTTCCAAATCAACCCAATCTTTTACCBCAAGCTGGCGAGATGCGAAAGTTTATGGGGAAA	450
Db	309	GWAAANAAAAAAMACCGAMRCCACCACCAWWATAAAAAATACGAATGCGTAAAWATAAA	368
Qy	451	AGGATAAAATGGTTTGATGGTGAGCAAGTAAATCACTAAAAATGCATTAATTCGAAGATACTT	510
Db	369	GAAGTATAATAG ---TGATFWCGTGAGACAAAAAAGGGCCANTAWMAWAATAWA	424
Qy	511	CGAAAGAATTAGAACACAAAATGTAAAAAGTATCTGCATGTGATAAAATATGGCCTTAGCCATT	570
Db	425	ACAAAAAAMWAAAACWAGTWATAAWATTATAAAAAWMAWMAWMAAACAAAAAWA	484
Qy	571	TCCAAATATCTAAATGTCTCAACTCAAGTAAAAA : : : :	618
Db	485	CAAAAAA : : : : : : :	532

RESULT 13	CNS017V2	1101 bp	DNA	linear	GSS 26-JUL-1999
LOCUS	Drosophila melanogaster	genome survey sequence sp6 end of BAC			
DEFINITION	BACN37L10 of DrosBAC library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL108536				
VERSION	ALL08536.1	GI:5628840			
KEYWORDS	GSS.				
SOURCE	fruit fly.				
ORGANISM	Drosophila melanogaster				
	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;				
	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;				
	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE	1 (bases 1 to 1101)				

Genoscope.
AUTHORS
TITLE
Direct Submission
JOURNAL
Submitted (23-JUL-1999) Genoscope - Centre National de Sequençage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT
Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
library (Dros BAC) was made by Alain Billaud at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC
project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelOBC11.

FEATURES	source	Location/Qualifiers			
		1. .li101			
		/organism="Drosophila melanogaster"			
		/plasmid="pbelOBAC11"			
		/db_xref="taxon:7227"			
		/clone_lib="DrosBAC"			
		/clone="BACN37110"			
		/note="end : SP6"			
BASE COUNT	268 a	174 c	114 g	366 t	179 others

Query Match 8.9%; Score 55.6; DB 12; Length 1101;
Best Local Similarity 40.9%; Pred. No. 5.1;
Matches 130; Conservative 33; Mismatches 155; Indels 0; Gaps 0;

QY 300 TACAGAACAAACTCAAGCCGAGTACATATCCTAACCAGAAAATGAAAAATCATATAAT 359
|| : || : | : || : || : || ||| |||| | || : ||
Db 1093 TATTWAWWWAATATWAHATWAWWAWATWATWTWTTWAAAAAAAAAAAAAAAAAAAA 1034

Qy	360	CAGGAAGGTAATTCGAGT	TTAATAACACTAACTTCCAAATCAACCCAATCTTTCACGC	419
		: : : : : : :		
Dd	1033	AAAAAAHAAWAAAAATTTTAAWWTAAAAAWATHATATAAAA	WWAAAAWATATATTTTTT	974
Qy	420	AAGCTGCCAGATGGGAAAGT	TATGGGAAAGGATAAAATGGTGTGCTGGAGCAGTA	479
		: : : :		
Dd	973	AAAYAAAYAAATAAWAAAAA	WAAAAWTAATAATATATAAAAAAAA	914
Qy	480	ATCACTAAAAATGACAATT	TGGAAGATCTTCGAAAGAAATTAGAAAAAAATGTAAAGT	539
		: : : :		
Dd	913	AAAAAAAAAAAAA	WAAAAAATAAAAAA	854
Qy	540	ATCTGCATTGATAATATG	CGCCITAGCCATTCCAAATATCTAAATGTCAACTCAGTA	599
		: : :		
Dd	853	AAAAATAAWAAATATATA	YAAAAA	794
Qy	600	AAAAAAAAAAAAAAAAAAAA	AAAAA	617
		:		
Dd	793	TAAWAAAAAAAAAAAAAA	AAAAA	776

RESULT 14	CNS008K5	CNS008K5	430 bp	DNA	linear	GSS 03-JUN-1999
LOCUS						
DEFINITION						Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR17M14 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION AL052179
 VERSION AL052179.1
 KEYWORDS GI:4933829
 SOURCE GSS:
 ORGANISM *Drosophila melanogaster*
 fruit fly.
 Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta; Pterygota; Neoptera; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE 1 (bases 1 to 430)

REFERENCE
AUTHORS
TITLE
JOURNAL

1 (bases 1 to 430)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the *Drosophila melanogaster* genome using these BACs. For further information please see <http://www.fruitfly.org> The BDGP *Drosophila melanogaster* BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPci-98 and was constructed by partial EcoRI digestion of *Drosophila* DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at <http://bacpac.med.buffalo.edu/drosophila.bac.htm>.

FEATURES	SOURCE
1. High Accuracy: The model achieves a high accuracy rate, consistently performing well across various datasets and tasks.	Model Performance Metrics
2. Scalability: The system is designed to scale efficiently, handling large volumes of data and complex computations.	System Architecture
3. Interpretability: The model's decisions are transparent and explainable, allowing users to understand the underlying reasoning.	Model Explainability
4. Robustness: The system is resilient to adversarial attacks and maintains performance in noisy or incomplete data environments.	Security and Reliability
5. Integration: The solution seamlessly integrates with existing infrastructure and third-party services.	System Integration

FEATURES	Source	Location/Qualifiers
		1..430
		/organism="Drosophila melanogaster"
		/accession="taxid:7227"
		/dbkey="taxid:7227"
		/clone="BAC17M14"
		/note="end : TET3"
BASE COUNT	331 a	27 c 18 g 10 t 44 others
ORIGIN		

```

Query Match      8.9%; Score 55.4; DB 12; Length 430;
Best Local Similarity 42.7%; Pred. No. 7.6;
Matches 146; Conservative 17; Mismatches 179; Indels 0; Gaps 0;

QY 276 ACTCAGAAGGTTACCAATCACTTCTTACAGAACAACACTCAAGCGACGTACATATCCTAAC 335
      |||||: ||| |||

```

```
Db 7 ACAGAGAAATGAANNACNNNNNNAACCCCAANNANACANACACMCAATCAAGA 66
QY 336 CAAGAAATGAAATCATATAATCAGGAAGGTAATTCGAGTTTATATACACATAATCTT 395
Db 67 AAAAAAAAAACACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAA 126
QY 396 CCAATCAACCCCAATCTTTCCAGCAAGCTGGCAGATCTGGAAGTTATGGGAAAGGAT 455
Db 127 AARAAACCAAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 186
QY 456 AAAATGGTTGATGTGAGCAAGTAATCACTAAAAATGACATATTATGAAGTACTTCCAAA 515
Db 187 ATATAAATTAAGGAAAAAARMAAARMAAARMAAARMAAARMAAARMAAARMAAARMA 246
QY 516 GAAATAGAACAAATGTAAGTATCTGCATGATAAATATGCGCTTACCCATTTCCAA 575
Db 247 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 306
QY 576 ATATCTAAATTCCACTCAAGTAAAAAARMAAARMAAARMAAARMAAARMAAARMA 617
Db 307 MAAAAAAAAAAGMCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 348

RESULT 15
CNSOLUPZ
LOCUS
DEFINITION
Tetraodon nigroviridis genome survey sequence PUC-Ori end of clone
197M17 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION
AL168128
VERSION
AL168128.1 GI:7806185
KEYWORDS
GSS; genome survey sequence.
SOURCE
Tetraodon nigroviridis.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
REFERENCE
1 (bases 1 to 581)
Roest-Crollius,H., Jallion,O., Dasilva,C., Fizames,C., Fisher,C.,
Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
Weissenbach,J.
Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
Unpublished
2 (bases 1 to 581)
Roest-Crollius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence
Unpublished
3 (bases 1 to 581)
Genoscope.
Direct Submission
Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.
FEATURES
source
Location/Qualifiers
1. .581
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="197M17"
/clone_lib="G"
/note="Genoscope sequence ID : COAG197AG09SP1-end :
PUC-Ori"
BASE COUNT 417 a 7 c 58 g 31 t 68 others
ORIGIN
Query Match 8.9%; Score 55.2; DB 12; Length 581;
Best Local Similarity 39.6%; Pred. No. 7.4;
```

```
Matches 141; Conservative 23; Mismatches 192; Indels 0; Gaps 0;
QY 262 AGAAGAGGATTGGACTCAGAAGCTTACCATAATCACTTCTTACAGAACAACTCAAGCCAG 321
Db 212 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 271
QY 322 CTACATATCTCAACCAAGAAATGAAATCATATAATCAAGGAAGGTAAATTCGAGTTTAA 381
Db 272 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 331
QY 382 ATACACTAAATCTTCCAAATCAACCCCAATCTTTCACGCAAGCTGGCAGATGTGGAAGTT 441
Db 332 AATGCAATAAWAWAWWWAAAAAATCAATCAATTAATTAATTAATTAATTAATTAATTA 391
QY 442 ATGGGGAAAAAGGATAAAATGTTGATGGTGAGCAAGTAATCACTTAAAAATGACATTTATG 501
Db 392 AAAAAAAAAAWATTAAGAAAGAAWARGAGRAAAAAATAAATAAAAAAAAAAAAAARR 451
QY 502 AAGATACTTCCAAAGAAATTTAGAAACAAATGTAAAGTATCTCGATTTGATAATATGCCC 561
Db 452 AAGGAATAWAAAAAARMAAARMAAARMAAARMAAARMAAARMAAARMAAARMAAARMA 511
QY 562 TTAGCCATTTCCAAATATCTTAAATTTGCAACTCAAGTAAAAAARMAAARMAAARMA 617
Db 512 TAAAAATAAAAAAATAAAAAAARMAAARMAAARMAAARMAAARMAAARMAAARMAAARMA 567

Search completed: November 2, 2002, 01:47:26
Job time : 1112.55 secs
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2002, 23:16:04 ; Search time 2065.61 Seconds
(without alignments)
7331.704 Million cell updates/sec

Title: us-09-881-556-3

Perfect score: 700

Sequence: 1 gaattggcagcagattttt.....acgcctctacagaaagc 700

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 21979536 seqs, 10817449327 residues

Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_NA_Main.*

1: /cgn2_6/ptodata/1/pna/PCTUS_COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US06_COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US07_COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US080_COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US081_COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US082_COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US083_COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US084_COMB.seq.*
9: /cgn2_6/ptodata/1/pna/US085_COMB.seq.*
10: /cgn2_6/ptodata/1/pna/US086_COMB.seq.*
11: /cgn2_6/ptodata/1/pna/US087_COMB.seq.*
12: /cgn2_6/ptodata/1/pna/US088_COMB.seq.*
13: /cgn2_6/ptodata/1/pna/US089_COMB.seq.*
14: /cgn2_6/ptodata/1/pna/US090_COMB.seq.*
15: /cgn2_6/ptodata/1/pna/US091_COMB.seq.*
16: /cgn2_6/ptodata/1/pna/US092_COMB.seq.*
17: /cgn2_6/ptodata/1/pna/US093_COMB.seq.*
18: /cgn2_6/ptodata/1/pna/US094_COMB.seq.*
19: /cgn2_6/ptodata/1/pna/US095A_COMB.seq.*
20: /cgn2_6/ptodata/1/pna/US095B_COMB.seq.*
21: /cgn2_6/ptodata/1/pna/US095C_COMB.seq.*
22: /cgn2_6/ptodata/1/pna/US095D_COMB.seq.*
23: /cgn2_6/ptodata/1/pna/US096A_COMB.seq.*
24: /cgn2_6/ptodata/1/pna/US096B_COMB.seq.*
25: /cgn2_6/ptodata/1/pna/US096C_COMB.seq.*
26: /cgn2_6/ptodata/1/pna/US096D_COMB.seq.*
27: /cgn2_6/ptodata/1/pna/US096E_COMB.seq.*
28: /cgn2_6/ptodata/1/pna/US097A_COMB.seq.*
29: /cgn2_6/ptodata/1/pna/US097B_COMB.seq.*
30: /cgn2_6/ptodata/1/pna/US097C_COMB.seq.*
31: /cgn2_6/ptodata/1/pna/US098A_COMB.seq.*
32: /cgn2_6/ptodata/1/pna/US098B_COMB.seq.*
33: /cgn2_6/ptodata/1/pna/US098C_COMB.seq.*
34: /cgn2_6/ptodata/1/pna/US099A_COMB.seq.*
35: /cgn2_6/ptodata/1/pna/US099B_COMB.seq.*
36: /cgn2_6/ptodata/1/pna/US099C_COMB.seq.*
37: /cgn2_6/ptodata/1/pna/US100_COMB.seq.*
38: /cgn2_6/ptodata/1/pna/US101_COMB.seq.*
39: /cgn2_6/ptodata/1/pna/US6000_COMB.seq.*
40: /cgn2_6/ptodata/1/pna/US6001_COMB.seq.*
41: /cgn2_6/ptodata/1/pna/US6002_COMB.seq.*
42: /cgn2_6/ptodata/1/pna/US6003_COMB.seq.*
43: /cgn2_6/ptodata/1/pna/US6004_COMB.seq.*

44: /cgn2_6/ptodata/1/pna/US6005_COMB.seq.*
45: /cgn2_6/ptodata/1/pna/US6006_COMB.seq.*
46: /cgn2_6/ptodata/1/pna/US6007_COMB.seq.*
47: /cgn2_6/ptodata/1/pna/US6008_COMB.seq.*
48: /cgn2_6/ptodata/1/pna/US6009_COMB.seq.*
49: /cgn2_6/ptodata/1/pna/US6010_COMB.seq.*
50: /cgn2_6/ptodata/1/pna/US6011_COMB.seq.*
51: /cgn2_6/ptodata/1/pna/US6012_COMB.seq.*
52: /cgn2_6/ptodata/1/pna/US6013_COMB.seq.*
53: /cgn2_6/ptodata/1/pna/US6014_COMB.seq.*
54: /cgn2_6/ptodata/1/pna/US6015_COMB.seq.*
55: /cgn2_6/ptodata/1/pna/US6016_COMB.seq.*
56: /cgn2_6/ptodata/1/pna/US6017_COMB.seq.*
57: /cgn2_6/ptodata/1/pna/US6018_COMB.seq.*
58: /cgn2_6/ptodata/1/pna/US6019_COMB.seq.*
59: /cgn2_6/ptodata/1/pna/US6020_COMB.seq.*
60: /cgn2_6/ptodata/1/pna/US6021_COMB.seq.*
61: /cgn2_6/ptodata/1/pna/US6022_COMB.seq.*
62: /cgn2_6/ptodata/1/pna/US6023_COMB.seq.*
63: /cgn2_6/ptodata/1/pna/US6024_COMB.seq.*
64: /cgn2_6/ptodata/1/pna/US6025_COMB.seq.*
65: /cgn2_6/ptodata/1/pna/US6026_COMB.seq.*
66: /cgn2_6/ptodata/1/pna/US6027_COMB.seq.*
67: /cgn2_6/ptodata/1/pna/US6028_COMB.seq.*
68: /cgn2_6/ptodata/1/pna/US6029_COMB.seq.*
69: /cgn2_6/ptodata/1/pna/US6030_COMB.seq.*
70: /cgn2_6/ptodata/1/pna/US6031_COMB.seq.*
71: /cgn2_6/ptodata/1/pna/US6032_COMB.seq.*
72: /cgn2_6/ptodata/1/pna/US6033_COMB.seq.*
73: /cgn2_6/ptodata/1/pna/US6034_COMB.seq.*
74: /cgn2_6/ptodata/1/pna/US6035_COMB.seq.*
75: /cgn2_6/ptodata/1/pna/US6036_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	700	100.0	700	33	US-09-881-556-3
2	54.8	7.8	342	33	US-09-865-419A-12796
3	53.6	7.7	481	31	US-09-823-301-7836
4	52.6	7.5	236	26	US-09-669-817A-17885
5	51	7.3	362	33	US-09-865-419A-5928
6	50.2	7.2	252	17	US-09-388-907-6637
7	50	7.1	445	26	US-09-666-355A-5285
8	50	7.1	537	37	US-10-021-323-7228
9	50	7.1	537	64	US-60-255-619-7228
10	49.2	7.0	4657	64	US-60-258-273-22
11	48.6	6.9	419	18	US-09-480-903-11234
12	48.6	6.9	419	36	US-09-960-352-11234
13	48.4	6.9	423	17	US-09-394-745-4413
14	48.4	6.9	423	22	US-09-565-306-68010
15	48.4	6.9	423	25	US-09-654-617-330054
16	48.4	6.9	423	27	US-09-684-016-330054
17	48.2	6.9	552	58	US-60-197-873-63174
18	48.2	6.9	1141	31	US-09-806-708A-22
19	48	6.9	412	58	US-60-197-872-74895
20	48	6.9	2000	33	US-09-887-274A-5263
21	47.8	6.8	445	19	US-09-521-640-237677
22	47.8	6.8	445	53	US-60-140-769-32426
23	47.8	6.8	541	33	US-09-865-439A-110751
24	47.6	6.8	605	7	US-08-358-918-36
25	47.6	6.8	832	24	US-09-621-976-2813
26	47.6	6.8	832	53	US-60-147-499-2813
27	46.6	6.7	399	29	US-09-758-109-12204
28	46.6	6.7	874	29	US-09-726-802-1504
29	46.6	6.7	978	19	US-09-521-640-274601
30	46.4	6.6	510	59	US-60-207-458-49190
31	46.4	6.6	399	64	US-60-253-652-10503
					Sequence 3, Appli
					Sequence 12796, A
					Sequence 7836, Ap
					Sequence 17885, A
					Sequence 5928, Ap
					Sequence 6627, Ap
					Sequence 5285, Ap
					Sequence 7228, Ap
					Sequence 7228, Ap
					Sequence 11234, A
					Sequence 11234, A
					Sequence 44113, A
					Sequence 68010, A
					Sequence 330054, A
					Sequence 330054, A
					Sequence 63174, A
					Sequence 22, Appli
					Sequence 74895, A
					Sequence 5263, Ap
					Sequence 237677, A
					Sequence 32426, A
					Sequence 110751, A
					Sequence 36, Appli
					Sequence 2813, Ap
					Sequence 12204, A
					Sequence 1504, Ap
					Sequence 274601, A
					Sequence 49190, A
					Sequence 10503, A

C 32 46.2 6.6 418 17 US-09-304-517A-259577 Sequence 259577,
C 33 46.2 6.6 418 36 US-09-985-678-259577 Sequence 259577,
C 34 46.2 6.6 446 17 US-09-371-146A-259449 Sequence 259449,
C 35 46.2 6.6 446 22 US-09-565-240-8353 Sequence 8353, Ap
C 36 46.2 6.6 446 25 US-09-654-617-67534 Sequence 67534, A
C 37 46.2 6.6 446 27 US-09-684-016-67534 Sequence 67534, A
C 38 46.2 6.6 451 58 US-60-197-872-53836 Sequence 53836, A
C 39 45.8 6.5 408 33 US-09-865-419A-1419 Sequence 1419, Ap
C 40 45.6 6.5 319 25 US-09-654-617-317618 Sequence 317618,
C 41 45.6 6.5 319 27 US-09-684-016-317618 Sequence 317618,
C 42 45.6 6.5 339 17 US-09-394-745-12853 Sequence 12853, A
C 43 45.6 6.5 339 22 US-09-565-306-33296 Sequence 33296, A
C 44 45.6 6.5 360 58 US-60-197-872-11831 Sequence 11831, A
C 45 45.6 6.5 565 31 US-09-824-130-4331 Sequence 4331, Ap

ALIGNMENTS

RESULT 1
US-09-881-556-3
; Sequence 3, Application US/09881556
; GENERAL INFORMATION:
; APPLICANT: Simonsen, J. Neil
; TITLE OF INVENTION: Cryptosporidium Parvum Antigens, Antibodies Thereof and
; TITLE OF INVENTION: Diagnostic and Therapeutic Compositions Thereof
; FILE REFERENCE: 9000-0054
; CURRENT APPLICATION NUMBER: US/09/881,556
; PRIOR FILING DATE: 2001-09-10
; PRIOR FILING DATE: 60/212,083
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 3
; LENGTH: 700
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:C. Parvum AG2
US-09-881-556-3

Query Match 100.0%; Score 700; DB 33; Length 700;
Best Local Similarity 100.0%; Pred. No. 9.6e-116; Indels 0; Gaps 0;
Matches 700; Conservative 0; Mismatches 0;
QY 1 GAATTCGGCAGAGATTTTCTTTTACCTATTTCAATTTAGTTCTTTGATTCAA 60
DB 1 GAATTCGGCAGAGATTTTCTTTTACCTATTTCAATTTAGTTCTTTGATTCAA 60
QY 61 ACAGTCAAAAGTCATTTTGTAAATCCAGATGGATCCGGAATTTTGAACACATTT 120
DB 61 ACAGTCAAAAGTCATTTTGTAAATCCAGATGGATCCGGAATTTTGAACACATTT 120
QY 121 CTACTAAATTCGAATTTAAATTTGAGTTGGCTTCAATAGTGGCTCGGAGGTG 180
DB 121 CTACTAAATTCGAATTTAAATTTGAGTTGGCTTCAATAGTGGCTCGGAGGTG 180
QY 181 ATATTTTATCCTTGATAGGAACACGCTCTTGAAGCTGTAAGTTATCAATCGCTTG 240
DB 181 ATATTTTATCCTTGATAGGAACACGCTCTTGAAGCTGTAAGTTATCAATCGCTTG 240
QY 241 TTTTCTATACAAAACATGTTTGAAGAAGATGAAGACATTTGCTTAAACCCCTTTAATC 300
DB 241 TTTTCTATACAAAACATGTTTGAAGAAGATGAAGACATTTGCTTAAACCCCTTTAATC 300
QY 301 GCGCTGAGAATAAATGACTTTTGGTCTGAGAAAGACTTAGCGCAACTCTCCAATCTT 360
DB 301 GCGCTGAGAATAAATGACTTTTGGTCTGAGAAAGACTTAGCGCAACTCTCCAATCTT 360
QY 361 CTAATTCGAATATTTATCTTTTCCCTTACATGGAATTAACATGCTTCTGGATATATCCAA 420
DB 361 CTAATTCGAATATTTATCTTTTCCCTTACATGGAATTAACATGCTTCTGGATATATCCAA 420

QY 421 TTAACACAAATTAATTAACACAAATTTCTCTTGAAGTTCCGGAGAAACTCAATCTCA 480
DB 421 TTAACACAAATTAATTAACACAAATTTCTCTTGAAGTTCCGGAGAAACTCAATCTCA 480
QY 481 CAATTGGATATTGGAGTATTATCGATGGATTTCTCTTCTTTAATTAACATGCGCTA 540
DB 481 CAATTGGATATTGGAGTATTATCGATGGATTTCTCTTCTTTAATTAACATGCGCTA 540
QY 541 TAAAGAAATGGCCACTTTGAATTAATCAAGAATCAAAATATTCAAAATGAATTAATGAAG 600
DB 541 TAAAGAAATGGCCACTTTGAATTAATCAAGAATCAAAATATTCAAAATGAATTAATGAAG 600
QY 601 CCATTAACATCAACAAATCCAGATCAGTGGGGAATACTTAACTCAGAACCAACACACA 660
DB 601 CCATTAACATCAACAAATCCAGATCAGTGGGGAATACTTAACTCAGAACCAACACACA 660
QY 661 AAGCCTCATCCAGTTGTTAGCCGCATCTTACAGAAAGC 700
DB 661 AAGCCTCATCCAGTTGTTAGCCGCATCTTACAGAAAGC 700

RESULT 2
US-09-865-419A-12796/C
; Sequence 12796, Application US/09865419A
; GENERAL INFORMATION:
; APPLICANT: Wu, Kunsheng
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(51935)B
; CURRENT APPLICATION NUMBER: US/09/865,419A
; PRIOR FILING DATE: 2001-05-29
; PRIOR FILING DATE: 60/208,063
; NUMBER OF SEQ ID NOS: 54020
; SEQ ID NO 12796
; LENGTH: 342
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3236-147-P1-N1-H5
US-09-865-419A-12796

Query Match 7.8%; Score 54.8; DB 33; Length 342;
Best Local Similarity 49.3%; Pred. No. 3.4;
Matches 143; Conservative 0; Mismatches 147; Indels 0; Gaps 0;
QY 373 ATTATCTTTTCCCTTACATGGAAATAACTGCATTTCTGGATATATTCCAAATTAACACAAATA 432
DB 294 ATTTATATTTTTTTTAAAAAAGTAACTTTTCTTTTAAATATAATAATGAATAATTTTTTATT 235
QY 433 AAATCAACAAATTTCTCTTGAAGTTCCGGAGAAACTCAATCTCCACAATTTGGATATT 492
DB 234 ATACAAATAAATATTTCTTTTAAATAATTTTAAATAATTTTAAATAATTTTATAT 175
QY 493 GGAGTATTATCGATGGATTTTCTTCTTTAATTAACATCGCGCTATAAAAGAAATG 552
DB 174 AATGATTTTTTTTATTTTATTTTTTTTTTAAAAAATAAATAAATAAATAAATAA 115
QY 553 GCCACTTGAATTAACAGAAATCAAAATATTCAAAATGAATTAAGAACTAACTCA 612
DB 114 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 55
QY 613 ACAATCCAGATCAGGTGGGAATTAACCTTAACCTCAGAACCAACCAACCAACCAACCA 662
DB 54 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 5

RESULT 3
US-09-823-301-7836/C
; Sequence 7836, Application US/09823301
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.

```

; APPLICANT: Macbeth, Kyle J.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.2085-001
; CURRENT APPLICATION NUMBER: US/09/823,301
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: US 60/193,421
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 10002
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7836
; LENGTH: 481
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(481)
; OTHER INFORMATION: n = A,T,C or G
US-09-823-301-7836

```

	Query Match	7.7%	Score 53.6;	DB 31;	Length 481;
	Best Local Similarity	43.1%	Pred. No. 5.9;		
	Matches 146;	Conservative 0;	Mismatches 193;	Indels 0;	Gaps 0;
Qy	320	TTTTGGTCTCAGAAAGACCTTAGCGACAACCTCCCAATCTCTAAATCTCGAATATTATCT	379		
Db	367	TTTTNNATNTTAAAAAANTATNGTTTTTTTTTTTAAAAATTTTTTTTACANGGGTTTT	308		
Qy	380	TTTCCTTTACATGGAAATAACTGCATCTTGGATATATCCCAATTAACACAAATAAAATCAA	439		
Db	307	TTTTTTTCAAAAAAATAAAAAAANCNNNTTTTTTTTTTGNNTTCCCCCNAAANNAATAA	248		
Qy	440	CAAAATTTCTCTTGAAGTTCCGGAGAAAACTCAATCTCCAATTTGGATATTTGGAGTAT	499		
Db	247	AAAATNTTTTTTTTTTTTTTTTTTANAAAAFANNTTTTTTAAANNTTTTNAANNGNTTAA	188		
Qy	500	TATCGATGGATTTTCTTCTCTTTTAAATTAACATCGCGCTATAAAGAAAAATGCCACTT	559		
Db	187	ATTNAAAAAANAATTCNCGNCTTTTTTTTTTTTCCCNANNTAAAAAATAAAAAAATAA	128		
Qy	560	GAATTAATCAAGAAATCAAAATATTCAAAATGAATTAATGAAGCCACTTAACTCAACAATC	619		
Db	127	AAAAANTAAAAAANNNAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAA	68		
Qy	620	CAGAAATCAGGTGGAAATAACTTTAACTCAGAACCAAAAAACA	658		
Db	67	AAAAAANAANNAAGNAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAA	29		

RESULT 4
US-09-669-817A-17885
; Sequence 17885, Application US/09669817A
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Ruen, Yijun G.
; APPLICANT: Wallick, C. Kevin
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(51469)B
; CURRENT APPLICATION NUMBER: US/09/669,817A
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US 60/156,951
; PRIOR FILING DATE: 1999-09-30
; NUMBER OF SEQ ID NOS: 43701
; SEQ ID NO 17885
; LENGTH: 236
; TYPE: DNA
; ORGANISM: *Oryza sativa*
; FEATURE:
; OTHER INFORMATION: Clone ID: uc-osflcyp022e09b2
US-09-669-817A-17885

Query Match 7.58; Score 52.6; DB 26; Length 236;

	Best Local Similarity	55.18;	Pred. No. 8;	Mismatches	Conservative	0;	Indels	84;	Gaps	0;
Qy	484	TTGGATATTGGGATATTATCGATGGATTTCTTCTCTTTAATTAAACATCGGCCTATAA	543							
Db	49	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGAGGAAAAATTTAAAATAAAA	108							
Qy	544	AAGAATAATGGCCACTTTGAATTAATCAAGAATCAAATAATTCAAAATGAAATTAATCAAGCCA	603							
Db	109	AAAAAAGGTAAAAAANAANAATAAANAANAANAATAATTTGANAANAANAANAANAANAANT	168							
Qy	604	CTAAACTCAACAATAATCAGATTCAGTGGGAATTAACITTAACTCAGAACCAAAAACACRAAG	663							
Db	169	ATAAAAAATAAAAAAATAATTTAAGAAATAAATAAAAAAGAGGAAAAAAAAAAAAA	228							
Qy	664	CCTCATC	670							
Db	229	TATAATC	235							

```

RESULT 5
US-09-865-419A-5928/c
: Sequence 5928, Application US/09865419A
: GENERAL INFORMATION:
: APPLICANT: Connor, Timothy W.
: APPLICANT: Wu, Kunsheng
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
: TITLE OF INVENTION: Plants
: FILE REFERENCE: 38-21(51935)B
: CURRENT APPLICATION NUMBER: US/09/865,419A
: CURRENT FILING DATE: 2001-05-29
: PRIOR APPLICATION NUMBER: 2001-05-29
: PRIOR FILING DATE: 2000-05-31
: NUMBER OF SEQ ID NOS: 54020
: SEQ ID NO 5928
: LENGTH: 362
: TYPE: DNA
: ORGANISM: Zea mays
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (1)..(362)
: OTHER INFORMATION: unsure at all n locations
: OTHER INFORMATION: Clone ID: LIB3236-069-p1-N1-G7
US-09-865-419A-5928

```

Query Match	7.3%	Score 51:	DB 33:	Length 362:
Best Local Similarity	47.2%:	Pred. No. 16:		
Matches 150:	Conservative 0:	Mismatches 168:	Indels 0:	Gaps 0:

Qy	345	ACAACTCTCCCAATCTTCTTAATTCTCGAATAATTTATCTTTTTCCITTTACATGGAATTAAGTCGATTT	404
Db	342	AAAAATAAAAAAATTTATTAATTAANAANATATTTATTTAAAATTTTAAAATATAAANAATTT	283
Qy	405	CTTGATATATTTCCAATTTAACACAATAAAAATCAACAAAAATTTCTCTTTGAAAGTTCCCGA	464
Db	282	TTTTTTTATATAATATAATAAATTTAAAAAANAANAATTTTTTTTTTTTTTAAAAAACAATTTAA	223
Qy	455	GAAACTCAATCTCCACAATTTGGATATATTCGATGGATTTCTCTCTCTTTA	524
Db	222	AATTTAAAAAANAANAANAANAANAATTTTTTTTAAAAAATTTCACTAAAAATTTTTTTTTTTTTT	163
Qy	525	ATTAACATGCCTTATAAAAAGAAATGCCACTTTGAATAATCAAGAATCAAAATATTCA	584
Db	162	TTTTTAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA	103
Qy	585	AAATGAAATTAATGAAGCCACTAAACTCAACAAATCCAGAAATCAGGTGGGAATTAAGTTAAC	644
Db	102	AAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA	43
Qy	645	TCAGAACCAAAACACAAA	662
Db	42	AAAAAAAAAAAAAAAAAAAA	25

RESULT 6

US-09-388-907-6627/c

; Sequence 6627, Application US/09388907

; GENERAL INFORMATION:

; APPLICANT: Searle Monsanto

; APPLICANT: Doshi, Parul D.

; APPLICANT: Dotson, Stanton B.

; APPLICANT: Chen, Jin

; APPLICANT: Ma, Xiao-Jun

; TITLE OF INVENTION: Nucleic acid molecules and other molecules associated with HS-5

; FILE REFERENCE: SO-3145

; CURRENT APPLICATION NUMBER: US/09/388,907

; EARLIER FILING DATE: 1999-09-02

; EARLIER FILING DATE: 1998/09/04

; NUMBER OF SEQ ID NOS: 7558

; SEQ ID NO 6627

; LENGTH: 252

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: Clone ID: LIB33-004-Q1-B1-C6

US-09-388-907-6627

Query Match

Best Local Similarity 7.2%; Score 50.2; DB 17; Length 252;

Matches 97; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 488 ATATTGGAGTATTATCGATGGGATTTCTTCTTAAATTAACATCGCGCTATAAAGA 547

DB 188 ATTTAAAAAAATTTTGTGAATTTTAAATTCAAAATTAATATTCCTGAAAAAAA 129

QY 548 AATGGCCACTTGAATCAAGAAATCAAAATATTCAAAATGAATCAAGCCACTAA 607

DB 128 AATTTTCCCTTTAAAGTTAAAAAATAAATAAATAAATAAATAAATAAATAA 69

QY 608 ACTCAACAATCCAGATCAGGTGGGAATTAACCTTAACCTCAGAACCAACACAAA 662

DB 68 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 14

RESULT 7

US-09-666-355A-5285

; Sequence 5285, Application US/09666355A

; APPLICANT: Fincher, Karen L.

; APPLICANT: La Rosa, Thomas J.

; APPLICANT: Xiao, Jinhua

; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With

; FILE REFERENCE: 38-21(51452)B

; CURRENT APPLICATION NUMBER: US/09/666,355A

; EARLIER FILING DATE: 2000-09-20

; EARLIER FILING DATE: 1999-09-21

; NUMBER OF SEQ ID NOS: 25978

; SEQ ID NO 5285

; LENGTH: 445

; TYPE: DNA

; ORGANISM: Gossypium hirsutum

; FEATURE:

; OTHER INFORMATION: Clone ID: uC-gsrnu33B169f10b1

US-09-666-355A-5285

Query Match

Best Local Similarity 7.1%; Score 50; DB 26; Length 445;

Matches 122; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 421 TTAACACAAATAAATCAACAAAATTTCTTTGAAAGTTCGCGAGAAACAAATCTCCA 480

DB 117 TTAATTTAGGTAAGAAAAATCAAAATGCTCAAAAGCATTGCTAGAGATTATTGAAGTCAAA 176

RESULT 9

US-60-255-619-7228

; Sequence 7228, Application US/60255619

QY 481 CAATTGGATATTGGAGTATTATCGATGGATTTCTTCTTAAATTAACATGCGCCCTA 540
DB 177 TGAATGGCTTTAGGAATTTGTAATGAAATTCATATCTATATCCAAAAAATAAATAA 236
QY 541 TAAAGAAAAATGCCCACTTGAATAATCAAGAAATCAAAATATTCAAAATGAATGAAG 600
DB 237 TAAAGAAAAATGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 296
QY 601 CCACTAACTCAACAAATCCAGAAATCAGGTGGGAATTAACCTCAGAACCAACACACA 660
DB 297 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 356
QY 661 AA 662
DB 357 AA 358

RESULT 8

US-10-021-323-7228

; Sequence 7228, Application US/10021323

; GENERAL INFORMATION:

; APPLICANT: Deikman, Jill

; APPLICANT: Feng, Paul C.C.

; APPLICANT: Fincher, Karen L.

; APPLICANT: Ziegler, Todd E.

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(52274)B

; CURRENT APPLICATION NUMBER: US/10/021,323

; PRIOR FILING DATE: 2001-12-12

; PRIOR APPLICATION NUMBER: US 60/255, 619

; NUMBER OF SEQ ID NOS: 17880

; SEQ ID NO 7228

; LENGTH: 537

; TYPE: DNA

; ORGANISM: Gossypium hirsutum

; FEATURE:

; OTHER INFORMATION: Clone ID: LIB3828-015-Q1-N6-E10

US-10-021-323-7228

Query Match

Best Local Similarity 7.1%; Score 50; DB 37; Length 537;

Matches 152; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY 16 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 75

DB 71 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 130

QY 76 TATTTGTTTTAAATCCAGATGATCCGGAATTTTGAAAAACATTTTCTACTAAATTCGAAA 135

DB 131 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 190

QY 136 TTAATTTTGAGCTTGGCTTGATAAATGTTAGTGTGCTCGAGGTGATATTTTATCCTTG 195

DB 191 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 250

QY 196 ATAGAAACACGCTCTTGAAGCTGTAAGTATTCAATCGCTTGTCTTCTTCTATACAAAA 255

DB 251 TTGAGAAAGACCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 310

QY 256 CATGTTTTGAAAGAAATGAAGCACATTTCTTTAAACCCCTTTAATCGCTCGAGATAAAA 315

DB 311 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCT 370

QY 316 TCACCTTTTGGTTCGAGAAAGA 337

DB 371 AAAAAATTTTCTTTTAAAGAAAA 392

```

; GENERAL INFORMATION:
; APPLICANT: Deikman, Jill
; APPLICANT: Feng, Paul C.C.
; APPLICANT: Fincher, Karen L.
; APPLICANT: Ziegler, Todd E.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(52274)A
; CURRENT APPLICATION NUMBER: US/60/255,619
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 17880
; SEQ ID NO 7228
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; OTHER INFORMATION: Clone ID: LIB3828-015-Q1-N6-E10
US-60-255-619-7228

```

Query Match	7.1%	Score 50;	DB 64;	Length 537;
Best Local Similarity	47.2%	Pred. No. 26;		
Matches 152;	Conservative	0;	Mismatches 170;	Indels 0;
Qy	16	TTTTTTTTTTTCTTTTACCTATTTCAAATTAGTTTCTTTTGATTCAAACGATGCAAGTCAT	75	
Db	71	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	130	
Qy	76	TATTTGTTTAAATCCAGATGGATCCGGAAATTTGAAAAACATTTCTACTAAATTCGAAA	135	
Db	131	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAAAAAAANAANAATTTAAAAAANAANAANA	190	
Qy	136	TTAAATTTGAGCTTGGCTTGATAAATGGTAGTGCCTCGGAGGTATTTTTATCTCTG	195	
Db	191	AAAAATAAAAACCAAGCTAAAAAANAACGGTTTCGAAAAAGAAATTTTTATTATATTAA	250	
Qy	196	ATAGAAACACGCTCTTGAAGCTCTAAGTTATTCAATCGCTTGTTGTTTCTATACAAAA	255	
Db	251	TTGAGANAAGAGCAAAAAAANAANAATTTTTTTGGGGGGGGGGGGTAAAAAANA	310	
Qy	256	CATCTTTTGAAGAAATGAAGCACATTGTCTTAAACCTTTTAAATCGGCTCAGAAATAAA	315	
Db	311	TTTTTTTTTGAATAGGAAAAAANAAGACAAAAAANAATTTTTTTTTTTTTTAAAAAANA	370	
Qy	316	TGACTTTTGGTCTGAGAAAGA	337	
Db	371	AAAAATTTTTTTTTTAAAGAAA	392	

```

RESULT 10
US-60-258-273-22
; Sequence 22, Application US/60258273
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN NUCLEAR HORMONE RECEPTOR
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN NUCLEAR
; TITLE OF INVENTION: HORMONE RECEPTOR PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL001043-PROV
; CURRENT APPLICATION NUMBER: US/60/258,273
; CURRENT FILING DATE: 2000-12-27
; NUMBER OF SEQ ID NOS: 312
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 46657
; TYPE: DNA
; ORGANISM: Human
US-60-258-273-22

```

	Query Match	7.08;	Score 49.2;	DB 64;	Length 4657;
	Best Local Similarity	49.6%;	Pred.No.71;		
	Matches 126; Conservative	0;	Mismatches 128;	Indels	0; Gaps 0;
Qy	355	AATCTTCATAATTGCAATATTATCTTTCCITTACATGAAGAATAAAGTGCATTTCTTGATATA	414		
Dd	1447	AAATATGTGAATTATAAATAACTTTTAATATAAAAATATGTAATTTATAATACTTTTAAAAAT	1506		

Q7		333	AACTCCTGCGATGATATACCTTCCCTACATGGGAATGCATTCCTGGATATA	414
Db		1447	AAATATGTAAATTATAAANTACTTTTAATTATAAANTATGTAATTATANAAT	1506

Qy	415	TTCNATTTACACAATAAAATCAACAAAATTTCTCTTGAAGTTCGCGAGANAACCAA	474
Dd	1507	ATCTAATTATAAAATATGTAAATTATAAACACTTTTAATTATAAAATATCTAATTATAAACA	1566
Qy	475	TCGCCACAATTGGGATATTGGAGTATTATCGATGGATTTTTCTCTCTCTTTTAATTAAACATG	534
Dd	1567	TTTTAATTATAAAAAATTTTAATTATAAACACTTTTAATTATAAAATATTTAATTATAAATA	1628
Qy	535	CGCCTATAAAGAANAATGGCCACTTCGAATAATCAAGAATCAAAATATTTCAAAATCAAAATA	594
Dd	1627	TTTTAATTATAAAAAATTTTAATTATAAAATATTTTAATTATAAAATATTTAATTATAAATA	1688
Qy	595	ATGAAGCCACTAAA	608
Dd	1687	TTTTTAATTATATAAAA	1700

RESULT 11
 US-09-480-902-11234/c
 ; Sequence 11234, Application US/09480902
 ; GENERAL INFORMATION:
 ; APPLICANT: Byatt, John C.
 ; APPLICANT: Mathalagan, Nagappan
 ; APPLICANT: Tao, Mengbing
 ; APPLICANT: Warren, Wesley C.
 ; TITLE OF INVENTION: Nucleic Acid and Other Molecules Associated with Lac
 ; TITLE OF INVENTION: Muscle and Fat Deposition
 ; FILE REFERENCE: 10298/1
 ; CURRENT APPLICATION NUMBER: US/09/480,902
 ; CURRENT FILING DATE: 2000-01-11
 ; EARLIER APPLICATION NUMBER: USSN 60/115,707
 ; EARLIER FILING DATE: 1999-01-12
 ; NUMBER OF SEQ ID NOS: 15112
 ; SEQ ID NO 11234
 ; LENGTH: 419
 ; TYPE: DNA
 ; ORGANISM: Bos taurus
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: 48-LIB3058-052-Q1-R1-D8
 US-09-480-902-11234

	Query Match	6.9%;	Score 48.6;	DB 18;	Length 419;	
	Best Local Similarity	46.2%;	Pred. No. 45;			
	Matches 162; Conservative	0;	Mismatches 189;	Indels	0;	Gaps
Qy	310 ATAAAAACACATTGGTTCCTCAGAAAGACTTTAGCGACAACCTCCCAATCTCTTAATCTG	369				
Dd						
	376 ATAAAATAATTTTTATANAATAATTTTTTAAAAAAAAATTTTTTAAAAATTTTTTATTAT	317				
Qy	370 AATATATCTTTTCCTTGACATGGGAATTAAGTCATCTTTGGATATATTCCAATTTAACACAA	429				
Dd						
	316 TATTTTATTTAAATATATAATTTTTTAAAAAATTTTTTAAAAAATTTAATTTTAAAAATTTAAAAA	257				
Qy	430 ATAAATCAACAAAAATTTCTCTCGAAGCTTCGCGAGAAGAACTCAATCTCCACAATTTGGAT	489				
Dd						
	256 ATATAATATAAAAAATTTTTTTTTTTATATATTTAATAAAATTTAATAATTTTAACTTTTTTAAA	197				
Qy	490 ATTGGAGTATTATCGATGGATTTCCTCTCTTTAATTTAAACATCGCGCTATTAAGAAGAA	549				
Dd						
	196 ATTTTTTTATTTATTTATTTTTTAAAAAANNAANNAANNAANNAANNAANNAANNAANTNAN	137				
Qy	550 ATGCCACTTCGAATCAATCAAGAACTCAAATATTCAAAAATGAAATTAATCAAGCCACTTAAC	609				
Dd						
	136 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	77				

RESULT 12
US-09-960-352-11234/c


```

US-09-654-617-330054/c
; Sequence 330054, Application US/09654617
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 38-21(15097)D
; CURRENT APPLICATION NUMBER: US/09/654,617
; CURRENT FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 463173
; SEQ ID NO 330054
; LENGTH: 423
; TYPE: DNA
; ORGANISM: Zea mays
US-09-654-617-330054

Query Match          6.9%; Score 48.4; DB 25; Length 423;
Best Local Similarity 47.8%; Prod. No. 49;
Matches 172; Conservative 0; Mismatches 186; Indels 2; Gaps 1;

Qy 305 TGA GAATAAAATGACTTTTGTTCTCGAAGAAAGACTTAGCGACAACCTCTCCAAATCTTCTAA 364
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 387 TTA AATAAATAAATAATTA A A A A A A A A A A A A A A A A A A A A A A A A A A A A
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 365 TTCTGGAATATTATCTTTTCCTTACATCGAATAACTGCATCTCTCGATATATATTCCAATAA 424
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 327 ATATAAATAA A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 425 CACAATAAATAATCAACAAAATTTCTCTTTGAAAGTTCCGGAGAAAACCTCAATCTCCACAAT 484
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 267 AAAAAATTATATAACAATAA A A A A A A A A A A A A A A A A A A A A A A A A A A
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 485 TGGATATTGGAGTATTATCGATGGATTTCTCTCTTTTAAATTAACATGCGCCTATAAA 544
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 207 AATATATAAATTATAA A A A A A A A A A A A A A A A A A A A A A A A A A A A A
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 545 AGAAATGGCCACTTGAATAA--TCAAGAATCAAAATATTCCAAATGAAATTAATGAAGCC 602
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 147 ATTAATAA A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 603 ACTAAACTCAACA A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 87 CCAAAACCCCA A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A A
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: November 2, 2002, 00:37:01
Job time : 2081.61 secs

THIS PAGE BLANK (u8p70)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2002, 22:55:48 ; Search time 271.958 Seconds
(without alignments)
9300.473 Million cell updates/sec

Title: US-09-881-556-3
Perfect score: 700
Sequence: 1 gaattggcagcagattttt.....accgcctctacagaaagc 700

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2893907 seqs, 1806667767 residues

Total number of hits satisfying chosen parameters: 5787814

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq.*
 - 2: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq.*
 - 3: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq.*
 - 4: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq.*
 - 5: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq.*
 - 6: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq.*
 - 7: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq.*
 - 8: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq.*
 - 9: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq.*
 - 10: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq.*
 - 11: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq.*
 - 12: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq.*
 - 13: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq.*
 - 14: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq.*
 - 15: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	48.4	6.9	7040	12	US-10-172-086-13
C 2	47.2	6.7	5501	13	US-10-221-613-204
C 3	46.2	6.6	639	9	US-09-814-353-6171
C 4	46.2	6.6	639	9	US-09-814-353-12449
C 5	46.2	6.6	639	9	US-09-814-353A-6171
C 6	46.2	6.6	639	9	US-09-814-353A-12449
C 7	46	6.6	539	9	US-09-814-353-5852
C 8	46	6.6	539	9	US-09-814-353-12133
C 9	46	6.6	539	9	US-09-814-353A-5852
C 10	46	6.6	539	9	US-09-814-353A-12133
C 11	45.6	6.5	752	9	US-09-814-353-18656
C 12	45.6	6.5	752	9	US-09-814-353A-18656
C 13	45.4	6.5	874	13	US-10-198-846-5980
C 14	45.4	6.5	5680	13	US-10-221-613-3
C 15	45.2	6.5	327	9	US-09-785-276A-8158
C 16	45.2	6.5	477	9	US-09-918-995-15842
C 17	45.2	6.5	18283	13	US-10-221-613-326
C 18	45	6.4	6394	13	US-10-221-613-239

Sequence 5831, Ap	19	44.8	6.4	500	9	US-09-814-353-5831	Sequence 5831, Ap
Sequence 12112, A	20	44.8	6.4	500	9	US-09-814-353-12112	Sequence 12112, A
Sequence 5831, Ap	21	44.8	6.4	500	9	US-09-814-353A-5831	Sequence 5831, Ap
Sequence 12112, A	22	44.8	6.4	500	9	US-09-814-353A-12112	Sequence 12112, A
Sequence 8283, Ap	C 23	44.6	6.4	513	9	US-09-785-276A-8283	Sequence 8283, Ap
Sequence 5993, Ap	C 24	44.6	6.4	631	9	US-09-814-353-5993	Sequence 5993, Ap
Sequence 12272, A	C 25	44.6	6.4	631	9	US-09-814-353-12272	Sequence 12272, A
Sequence 5993, Ap	C 26	44.6	6.4	631	9	US-09-814-353A-5993	Sequence 5993, Ap
Sequence 12272, A	C 27	44.6	6.4	631	9	US-09-814-353A-12272	Sequence 12272, A
Sequence 7, Appl1	C 28	44.4	6.3	1799	12	US-10-155-533-7	Sequence 7, Appl1
Sequence 1, Appl1	C 29	44.4	6.3	2017	12	US-10-155-533-3	Sequence 1, Appl1
Sequence 17543, A	C 30	44.4	6.3	5314	12	US-10-155-533-1	Sequence 17543, A
Sequence 17543, A	C 31	44	6.3	352	9	US-09-814-353-17543	Sequence 17543, A
Sequence 323, App	C 32	44	6.3	392	9	US-09-814-353A-17543	Sequence 323, App
Sequence 1510, Ap	C 33	44	6.3	7131	13	US-10-221-613-323	Sequence 1510, Ap
Sequence 7598, Ap	C 34	43.8	6.2	359	9	US-09-789-189-1510	Sequence 7598, Ap
Sequence 11962, A	C 35	43.6	6.2	478	9	US-09-785-276A-7598	Sequence 11962, A
Sequence 914, App	C 36	43.4	6.2	491	9	US-09-918-995-11962	Sequence 914, App
Sequence 930, App	C 37	43.2	6.2	14809	2	PCT-US02-09188-914	Sequence 930, App
Sequence 6325, App	C 38	43.2	6.2	14809	2	PCT-US02-09370-950	Sequence 6325, App
Sequence 3, Appl1	C 39	43.2	6.2	14809	12	US-10-105-299-6325	Sequence 3, Appl1
Sequence 33, Appl1	C 40	43.2	6.2	53332	13	US-10-224-562-3	Sequence 33, Appl1
Sequence 55615, A	C 41	43	6.1	396	13	US-10-212-677-33	Sequence 55615, A
Sequence 113, App	C 42	42.8	6.1	512	9	US-09-785-276A-55615	Sequence 113, App
Sequence 413, App	C 43	42.6	6.1	9760	13	US-10-221-613-113	Sequence 413, App
Sequence 64, Appl1	C 44	42.4	6.1	5241	13	US-10-221-613-413	Sequence 64, Appl1
	C 45	42.4	6.1	6298	13	US-10-239-676-64	

ALIGNMENTS

RESULT 1

US-10-172-086-13/c

- Sequence 13, Application US/10172086
- GENERAL INFORMATION:
- APPLICANT: Epigenomics AG
- TITLE OF INVENTION: Method and nucleic acids for the differentiation of prostate tumors
- FILE REFERENCE:
- CURRENT APPLICATION NUMBER: US/10/172,086
- CURRENT FILING DATE: 2002-06-13
- NUMBER OF SEQ ID NOS: 116
- SEQ ID NO 13
- LENGTH: 7040
- TYPE: DNA
- ORGANISM: Artificial Sequence
- FEATURE:
- OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

Query Match 6.9%; Score 48.4; DB 12; Length 7040;

Best Local Similarity 52.9%; Pred. No. 2;

Matches 127; Conservative 0; Mismatches 111; Indels 2; Gaps 1;

Qy	423	AACACAAATAAATCACACAAATTTCTCTGAAAGTTCGGAGGANAACCTCAATCTCCACA	482
Db	1830	AAAACATATTTTAAATAATATTTAAATAATATCTATCAAACTAAATAAATCTCG	1771
Qy	483	ATTGGATTTGGAGTATATCGATGGATTTCTCTCTTTTAATTAACATGCGCTATA	542
Db	1770	TTTTCATATTAACATACATACAAATTTAAATATTTTCTACTCTTAATTTTACGATTTAA	1711
Qy	543	AAAGAAATGCCACTTGAATAATCAAGAAATCAAAATTAATCAAAATGAATGAAGCC	602
Db	1710	AAATTAATAATCTTACTATATATATATACACATATATATAAACGTAATATAAATTT	1651
Qy	603	ACTAACTCAACAAATCCAGAAATCAGGTGGGAATTAACCTAGAACCAACAAACAAA	662
Db	1650	A--ACACTAATCTAATCAATACATATATACCTATATATATATATATATATATATATAT	1593

RESULT 2

US-10-221-613-204/c
; Sequence 204, Application US/10221613

; GENERAL INFORMATION:

; APPLICANT: OLEK, Alexander

; APPLICANT: PIEPENBROCK, Christian

; APPLICANT: BERLIN, Kurt

; TITLE OF INVENTION: Diagnosis of Diseases Associated with Cell Cycle

; FILE REFERENCE: 5013.1004

; CURRENT APPLICATION NUMBER: US/10/221,613

; CURRENT FILING DATE: 2002-09-13

; PRIOR APPLICATION NUMBER: PCT/EP01/02945

; DE 10013847.00

; DE 10019058.8

; DE 10019173.8

; DE 10032529.7

; DE 10043826.1

; PRIOR FILING DATE: 2001-03-15

; 2000-03-15

; 2000-04-06

; 2000-04-07

; 2000-06-30

; 2000-09-01

; NUMBER OF SEQ ID NOS: 428

; SEQ ID NO 204

; LENGTH: 5501

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-221-613-204

Query Match 6.7%; Score 47.2; DB 13; Length 5501;

Best Local Similarity 46.6%; Pred. No. 3.3;

Matches 151; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 334 AAGACTTAGCGAACTCCCAATCTCTAATTCGAAATATTATCTTTTCTTACATGA 393

DB 2933 AATATATAAATAAACAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2874

QY 394 ATAATCGAATCTTGGATATATTCGAATTAACACAAATAAATAAATAAATAAATAA 453

DB 2873 TTAATAATATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2814

QY 454 AAGATCCGAGAACTCAATCTCCCAATGGATATGGAGTATATCGATGGATTTT 513

DB 2813 AATATCTTTAAACATTTTATCTTTAAATACCTATTATAAATAAATAAATAAATAA 2754

QY 514 CTTCTCTTTAATAAATGCGCTATATAAAGAAATGCGCACTTGGAATATCAAGAT 573

DB 2753 ACAACCTATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2694

QY 574 CAAATATTCAAATGAATAATGAAGCCACTTAACTCAACAAATCCAGAAATCAGGTGG 633

DB 2693 AAAAACTTCAAACTAAACATACCATTTTAAACAATAAATAAATAAATAAATAAATAA 2634

QY 634 ATAATCTTAACTCAGACCAAAAC 657

DB 2633 AATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 2610

RESULT 3

US-09-814-353-6171/c

; Sequence 6171, Application US/09814353

; GENERAL INFORMATION:

; APPLICANT: Lee, John

; APPLICANT: Thompson, Pamela

; APPLICANT: Lillie, James

; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR

; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND

; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER

; FILE REFERENCE: MRI-006B

; CURRENT APPLICATION NUMBER: US/09/814,353

; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: US 60/191,031

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: US 60/207,124

; PRIOR FILING DATE: 2000-05-25

; PRIOR APPLICATION NUMBER: US 60/211,940

; PRIOR FILING DATE: 2000-06-15

; PRIOR APPLICATION NUMBER: US 60/216,820

; PRIOR FILING DATE: 2000-07-07

; PRIOR APPLICATION NUMBER: US 60/220,661

; PRIOR FILING DATE: 2000-07-25

; PRIOR APPLICATION NUMBER: US 60/257,672

; PRIOR FILING DATE: 2000-12-21

; NUMBER OF SEQ ID NOS: 22037

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6171

; LENGTH: 639

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: 37, 127, 128, 129, 130, 159, 172, 181, 182, 187, 188, 190,

; LOCATION: 217, 218, 219, 228, 229, 230, 232, 233, 234, 239, 240, 253,

; LOCATION: 272, 276, 299, 300, 307, 314, 323, 324, 325, 328, 330, 359,

; LOCATION: 363, 364, 368, 383, 388, 423, 424, 428, 429, 430, 439

; OTHER INFORMATION: n = A,T,C or G

; NAME/KEY: misc_feature

; LOCATION: 444, 447, 452, 458, 463, 464, 470, 472, 474, 481, 482, 484,

; LOCATION: 485, 494, 508, 512, 516, 530, 533, 535, 536, 544, 545, 548,

; LOCATION: 553, 559, 564, 565, 567, 570, 571, 583, 587, 590, 597, 599,

; LOCATION: 608, 610, 613, 614, 623, 624, 634

; OTHER INFORMATION: n = A,T,C or G

US-09-814-353-6171

Query Match

Best Local Similarity 6.6%; Score 46.2; DB 9; Length 639;

Matches 186; Conservative 0; Mismatches 298; Indels 0; Gaps 0;

QY 175 GAGGTGATATTTTATCTCTGATAGGAACACGCTCTTGAAGCTGTAGCTTATCAATCG 234

DB 534 GNGNGTTTTTTTATTTGNTTNGGANAATAAATAAAGGGNTTTATAAANNNTAATTT 475

QY 235 CTTGTGTTTCTATACAAAACATGTTTGAAGAAGTGAAGCAGTCTCTTAAACCT 294

DB 474 NTNGNGGGGNNNTTAANAACCCCTCTNTTATAAANAATAAANAAGGNNNTTTTTTT 415

QY 295 TTAATCGCGCTGAGATATAAATGACTTTTGGTCTGAGAAAGACTTAGGCAACTCTCC 354

DB 414 TTTTNTTTTATTAATAAANAATAAANAAGGNTTTTAAANNNTAATTTT 355

QY 355 AATCTTCTAATTTCTGAATATATCTTTTCTTACATGGAATCACTTCTTGGATATA 414

DB 354 TTTTTCCTCAAAAAATAATTTTNTNGNCCAAAAAATTTTNTCCCTTTNNAAA 295

QY 415 TTCCAATTAACAAATAAATCAACAATAATTTCTTGTGAAGTTCGAGAGAACTCAA 474

DB 294 AAAAAATAAATAATTTTNTGCAAAAAATTTTNTTAAAAAATAAANAATTTT 235

QY 475 TCTCCACAATTTGGATATTATCGATGATTTTCTTCTTCTTCTTCTTCTTCTTCT 534

DB 234 NNNGNNNTTTTNTTNNNGGGGAAAAAATTTTNTTNTTNTTNTTNTTNTTNTTNT 175

QY 535 CGCCTATAAAGAAAAATGCCACTTTGAATTAATCAAGAATCAAAATATTCAAAATGAAATA 594

DB 174 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 115

QY 595 ATGAGCCACTTAACCTCAACAATCCAGATCAGTGGGAATACTTAACTCAGAACCAA 654

DB 114 AAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 55

QY 655 AACA 658

DB 54 AAAA 51

```
RESULT 4
; US-09-814-353-12449/c
; Sequence 12449, Application US/09814353
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12449
; LENGTH: 639
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 37, 127, 128, 129, 130, 159, 172, 181, 182, 187, 188, 190,
; LOCATION: 217, 218, 219, 228, 229, 230, 232, 233, 234, 239, 240, 253,
; LOCATION: 272, 276, 299, 300, 314, 323, 324, 325, 328, 330, 359,
; LOCATION: 363, 364, 368, 383, 388, 423, 424, 428, 429, 430, 439
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 444, 447, 452, 458, 463, 464, 470, 472, 474, 481, 482, 484,
; LOCATION: 485, 494, 508, 512, 516, 530, 533, 535, 536, 544, 545, 548,
; LOCATION: 553, 559, 564, 565, 567, 570, 571, 583, 587, 590, 597, 599,
; LOCATION: 608, 610, 613, 614, 619, 623, 624, 634
; OTHER INFORMATION: n = A,T,C or G
; US-09-814-353-12449

Query Match          6.68; Score 46.2; DB 9; Length 639;
Best Local Similarity 38.4%; Pred. No. 4.1;
Matches 186; Conservative 0; Mismatches 298; Indels 0; Gaps 0;

QY 175 GAGGTGATATTTTATCTTCATAGGAACACGCTCTTGAAGCTGTAACTTATTCATCG 234
DB 534 GNGNGTGTATTTTATTTGTTTNGGANAAGGGGNTTTTATAAANNNTAATTT 475

QY 235 CTTCGTGTTTCTATACAAAACATGTTTGAAGAAATGAAGCACATGTCTTAAACCC 294
DB 474 NTNGGGGGNNTTAAANAACCNCCTCTTNTAANAAGAAAGGNTTTTATTTT 415

QY 295 TTAATCCGCTGAGAATAAATGACTTTTGGTCTCGAAGAGACTTACCGCAACTCC 354
DB 414 TTTTGTGTTTATTTATTAAGAAAGGNTTTTAAANNNTAATTTT 355

QY 355 AATCTTCTAATTCGAATATATCTTTTCTACATGAATACTGCTTCTTGATATA 414
DB 354 TTTTTCCTCAAAAAAATTTTNTNGNNAACCAAAAAATTTTTCCTTTNAAAA 295

QY 415 TTCCAATTAACAAATAAATCAACAAAATTTCTCTTGAAGTTCCGGAGAAACTCAA 474
DB 294 AAAAAAATAATTTTNTGTCNAAAAATTTTNTTAAAAAATAAANNNTTTT 235
```

```
QY 475 TCTCCCAAATGGATATTTGGAGTATTTATCGATGATTTTCTTCTCTTAAATTAACATG 534
DB 234 NNNGNNNTTTTNTNNNGGGGAAAAAATTTTNTNTNTGAGNNAAAAA 175

QY 535 CGCCTATATAAGAAATGGCACCTTGAATATCAAGATCAAAATATTCAAAATGAATA 594
DB 174 AANAATATAAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 115

QY 595 ATGAAGCCACTAAACTCAACAAATCCAGAATCCAGGTGGGAATAAATTAAGTCAAGAACCA 654
DB 114 AAAAAAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANAANA 55

QY 655 AACA 658
DB 54 AAAA 51

RESULT 5
; US-09-814-353A-6171/c
; Sequence 6171, Application US/09814353A
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6171
; LENGTH: 639
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 37, 127, 128, 129, 130, 159, 172, 181, 182, 187, 188, 190,
; LOCATION: 217, 218, 219, 228, 229, 230, 232, 233, 234, 239, 240, 253,
; LOCATION: 272, 276, 299, 300, 307, 314, 323, 324, 325, 328, 330, 359,
; LOCATION: 363, 364, 368, 383, 388, 423, 424, 428, 429, 430, 439
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 444, 447, 452, 458, 463, 464, 470, 472, 474, 481, 482, 484,
; LOCATION: 485, 494, 508, 512, 516, 530, 533, 535, 536, 544, 545, 548,
; LOCATION: 553, 559, 564, 565, 567, 570, 571, 583, 587, 590, 597, 599,
; LOCATION: 608, 610, 613, 619, 623, 624, 634
; OTHER INFORMATION: n = A,T,C or G
; US-09-814-353A-6171

Query Match          6.68; Score 46.2; DB 9; Length 639;
Best Local Similarity 38.4%; Pred. No. 4.1;
Matches 186; Conservative 0; Mismatches 298; Indels 0; Gaps 0;

QY 175 GAGGTGATATTTTATCTTCATAGGAACACGCTCTTGAAGCTGTAACTTATTCATCG 234
DB 534 GNGNGTGTATTTTATTTGTTTNGGANAAGGGGNTTTTATAAANNNTAATTT 475

QY 235 CTTCGTGTTTCTATACAAAACATGTTTGAAGAAATGAAGCACATGTCTTAAACCC 294
```

Db 474 NTNGGGGNNNTTAAACACCCCTCNTTATTAATAAAAAAANNNGGNNTTTTTTT 415
QY 295 TTAATCGCGCTGAGAAATAAAATGACTTTTGGTTCTGAGAAAGACTTAGCGACAACCTCTCC 354
Db 414 TTTTCTTTTATTAATAAAAAAANCCCNCCGCCCGGGTTTNAANNAATNTTTT 355
QY 355 ATCTCTCTAATCTGAATATATCTTTCTTACATGGAATCACTGCAATCTTGGATATA 414
Db 354 TTTTCCCAAAAAAATTTTNGGNNCCAAAAAATTTTNNCCCTTNNAAAA 295
QY 415 TTCCCAATTAACAATAAAATCAACAAATTTCTCTGAAAGTTCCGGAGAAACTCAA 474
Db 294 AAAAAAATAAATTTTNGGNNCCAAAAAATTTTNNCCCTTNNAAAA 235
QY 475 TCTCCACAATGGATATGAGTATATCGATGGATTTCTTCTTTAAATAACATG 534
Db 234 NNNNGNNNTTTTNNNGGGGAAAAAATTTTNTTNTGAGNNAAAAA 175
QY 535 CGCCTATAAAGAAATGCCACTTGAATATCAAGATCAAAATATCAAAATGAATA 594
Db 174 AANAATAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 115
QY 595 ATGAAGCCACTAACTCAACAATCCAGATCAGGTGGGAATACTTAACCTCAGAACCA 654
Db 114 AAAAAAATAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 55
QY 655 AACA 658
Db 54 AAAA 51

RESULT 6

US-09-814-353A-12449/c
; Sequence 12449, Application US/09814353A
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12449
; LENGTH: 639
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 37, 127, 128, 129, 130, 159, 172, 181, 182, 187, 188, 190,
; LOCATION: 217, 218, 219, 228, 229, 230, 232, 233, 234, 239, 240, 253,
; LOCATION: 272, 276, 299, 300, 307, 314, 323, 324, 325, 328, 330, 359,
; LOCATION: 363, 364, 368, 383, 388, 423, 424, 428, 429, 430, 439
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 444, 447, 452, 458, 463, 464, 470, 472, 474, 481, 482, 484,

; LOCATION: 485, 494, 508, 512, 516, 530, 533, 535, 536, 544, 545, 548,
; LOCATION: 553, 559, 564, 565, 567, 570, 571, 583, 587, 590, 597, 599,
; LOCATION: 608, 610, 613, 614, 619, 623, 624, 634
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353A-12449
Query Match 6.6%; Score 46.2; DB 9; Length 639;
Best Local Similarity 38.4%; Pred. No. 4.1;
Matches 186; Conservative 0; Mismatches 298; Indels 0; Gaps 0;
QY 175 GAGGTGATATTTTATCTCTTGATAGGAACACGCTCTTGAAGCTGTAACTTATCAATCG 234
Db 534 GNGGNGTTTTTTTATTTGNTTTNGGANAATAAAAAAGGGGNTTTATAAANNNAATTTT 475
QY 235 CTGTGTGTTTCTATCAAAAAACATGTTTGAAGAATGAAGCACATGCTCTTAAACCT 294
Db 474 NTNGGGGGGNNNTTAANAAACCCCTCNTTATTAANAATAAAAAAANNNGGNNTTTTTT 415
QY 295 TTAATCGCGCTGAGAAATAAAATGACTTTTGGTTCTGAGAAAGACTTAGCGACAACCTCC 354
Db 414 TTTTCTTTTATTAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 355
QY 355 ATCTCTCTAATCTGAATATATCTTTCTTACATGGAATCACTTCTTGGATATA 414
Db 354 TTTTCCCAAAAAAATTTTNTGNNCCAAAAAATTTTNNCCCTTNNAAAA 295
QY 415 TTCCCAATTAACAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 474
Db 294 AAAAAAATAAATTTTNTGNNCCAAAAAATTTTNTGNNCCAAAAAATAAATAAATAAATAA 235
QY 475 TCTCCACAATGGATATGAGTATATCGATGGATTTCTTCTTTAAATAACATG 534
Db 234 NNNNGNNNTTTTNTGNNCCAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 175
QY 535 CGCCTATAAAGAAATGCCACTTGAATATCAAGATCAAAATATCAAAATGAATA 594
Db 174 AANAATAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 115
QY 595 ATGAAGCCACTAACTCAACAATCCAGATCAGGTGGGAATACTTAACCTCAGAACCA 654
Db 114 AAAAAAATAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 55
QY 655 AACA 658
Db 54 AAAA 51

RESULT 7

US-09-814-353-5852/c
; Sequence 5852, Application US/09814353
; GENERAL INFORMATION:
; APPLICANT: Lee, John
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037


```
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5852
; LENGTH: 539
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 74, 82, 152, 171, 177, 190, 195, 198, 207, 210, 215, 216,
; LOCATION: 222, 223, 224, 228, 230, 232, 241, 242, 245, 253, 254, 273,
; LOCATION: 293, 294, 296, 300, 302, 310, 312, 317, 319, 321, 325, 331,
; LOCATION: 333, 335, 336, 346, 348, 352, 369, 373, 375, 383, 388
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 393, 394, 403, 404, 409, 423, 425, 428, 443, 446, 447, 449,
; LOCATION: 451, 455, 465, 468, 469, 472, 474, 475, 478, 481, 488, 490,
; LOCATION: 491, 495, 496, 504, 513, 515, 518, 525, 526
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353A-5852

Query Match          6.6%; Score 46; DB 9; Length 539;
Best Local Similarity 41.7%; Pred. No. 4.3;
Matches 130; Conservative 0; Mismatches 182; Indels 0; Gaps 0;

QY 357 TCTTCTAAATCTCGAATATATCTTTTCCCTACATGGAATGAATCTTGGATATATT 416
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 TTTTITTTTGGGNANTTTTTTTTTTNNCNTNCAANAATNNANATTTNANAATATTN 302

QY 417 CCAATTAACAAATAAATCAACAAATTTCTCTTGAAGTTCGGAGAAAACCTCAATC 476
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 ANAATNTNTNGGNNNTAGTCNNTTTTTTCATCAAAAANNTTATCCNCCN 242

QY 477 TCCCAATTTGGATATTTGGAGTATATCGATGATTTCTCTTCTTAAATGAACATGCG 536
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 NAAAAATTTNTNGGNNNTAGTCNNTTTTTTCATCAAAAANNTTATCCNCCN 182

QY 537 CCTATAAAGAAATAGCCACTTGAATATCAAGAAATCAAAATATCAAAATGAATAAT 596
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 TTTANAAAAANTTTTTTTTCAAAAAANAANAATAAAAAAATAAAAAAATAAAAAA 122

QY 597 GAAGCCACTAACTCAACAAATCCAGATCAGTGGGAATACTTAACCTCAGAACCAAA 656
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 62

QY 657 CACAAGCCTCA 668
   ||||| |||||
Db 61 AAAAAAACCCA 50

RESULT 10
US-09-814-353A-12133/c
; Sequence 12133, Application US/09814353A
; GENERAL INFORMATION:
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
```

```
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12133
; LENGTH: 539
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 74, 82, 152, 171, 177, 190, 195, 198, 207, 210, 215, 216,
; LOCATION: 222, 223, 224, 228, 230, 232, 241, 242, 245, 253, 254, 273,
; LOCATION: 293, 294, 296, 300, 302, 310, 312, 317, 319, 321, 325, 331,
; LOCATION: 333, 335, 336, 346, 348, 352, 369, 373, 375, 383, 388
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 393, 394, 403, 404, 409, 423, 425, 428, 443, 446, 447, 449,
; LOCATION: 451, 455, 465, 468, 469, 472, 474, 475, 478, 481, 488, 490,
; LOCATION: 491, 495, 496, 504, 513, 515, 518, 525, 526
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353A-12133

Query Match          6.6%; Score 46; DB 9; Length 539;
Best Local Similarity 41.7%; Pred. No. 4.3;
Matches 130; Conservative 0; Mismatches 182; Indels 0; Gaps 0;

QY 357 TCTTCTAAATCTCGAATATATCTTTTCCCTACATGGAATGAATCTTGGATATATT 416
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 TTTTITTTTGGGNANTTTTTTTTTTNNCNTNCAANAATNNANATTTNANAATATTN 302

QY 417 CCAATTAACAAATAAATCAACAAATTTCTCTTGAAGTTCGGAGAAAACCTCAATC 476
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 ANAATNTNTNGGNNNTAGTCNNTTTTTTCATCAAAAANNTTATCCNCCN 242

QY 477 TCCCAATTTGGATATTTGGAGTATATCGATGATTTCTCTTCTTAAATGAACATGCG 536
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 241 NAAAAATTTNTNGGNNNTAGTCNNTTTTTTCATCAAAAANNTTATCCNCCN 182

QY 537 CCTATAAAGAAATAGCCACTTGAATATCAAGAAATCAAAATATCAAAATGAATAAT 596
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 181 TTTANAAAAANTTTTTTTTCAAAAAANAANAATAAAAAAATAAAAAAATAAAAAA 122

QY 597 GAAGCCACTAACTCAACAAATCCAGATCAGTGGGAATACTTAACCTCAGAACCAAA 656
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 121 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 62

QY 657 CACAAGCCTCA 668
   ||||| |||||
Db 61 AAAAAAACCCA 50

RESULT 11
US-09-814-353-18656/c
; Sequence 18656, Application US/09814353
; GENERAL INFORMATION:
; APPLICANT: Thompson, Pamela
; APPLICANT: Lillie, James
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF OVARIAN CANCER
; FILE REFERENCE: MRI-006B
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
```


; LOCATION: 657, 661, 667, 681, 682, 684, 689, 704, 707, 724, 742, 744,
; LOCATION: 751, 753, 758, 773, 774, 781, 782, 783, 784, 786, 798
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 800, 810, 812, 813, 818, 819, 828, 829, 840, 842, 850, 861,
; LOCATION: 869, 873, 874
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-5980

Query Match 6.5%; Score 45.4; DB 13; Length 874;
Best Local Similarity 39.5%; Pred. No. 6.1;
Matches 103; Conservative 0; Mismatches 158; Indels 0; Gaps 0;

QY 16 TTTTTCCTTTTCTTACCTATTCAATAGTTCTTCTTGAATCAACGATGCAAAAGTCAT 75
DB 114 TTTTTCCTTTTCTTACCTATTCAATAGTTCTTCTTGAATCAACGATGCAAAAGTCAT 173
QY 76 TATTCCTTTTAAATCCAGATCGGAAATTTGAAAAACATTTCTACTAAATTCGAAA 135
DB 174 TTTTTCCTTTTAAANAANAAGGNTTNTTTTNNAAAAAANNNNGCNCNANTCNA 233
QY 136 TTAATTTGAGCTTGCTTGATAAATGGTAGTTGGCTCGGAGGTGATATTTTATCCTTG 195
DB 234 AAAAAATNNNNNNCCCCCCCCNTTNCARAGGNNNANNNNNNGGNNNG 293
QY 196 ATAGAAACACGCTCTGAGCTGTAGTATTCAATCGCTGTGTCTATACAAAAA 255
DB 294 NNNGGNTTNTTTTNNAAAAAANAATNGTCNTCCCCCNNTTTTTTTTTTNNAAA 353
QY 256 CATGTTTGAAGAATGAAG 276
DB 354 NCCCCNTTNNAAAAAANAGG 374

RESULT 14
US-10-221-613-3
; Sequence 3, Application US/10221613
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Cell Cycle
; FILE REFERENCE: 5013.1004
; CURRENT APPLICATION NUMBER: US/10/221,613
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: PCT/EP01/02945
; DE 10013847.00
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-03-15
; 2000-03-15
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 428
; SEQ ID NO 3
; LENGTH: 5680
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-613-3

Query Match 6.5%; Score 45.4; DB 13; Length 5680;
Best Local Similarity 49.0%; Pred. No. 7.7;
Matches 121; Conservative 0; Mismatches 126; Indels 0; Gaps 0;

QY 17 TTTTTCCTTTTCTTACCTATTCAATAGTTCTTCTTGAATCAACGATGCAAAAGTCAT 76
DB 114 TTTTTCCTTTTCTTACCTATTCAATAGTTCTTCTTGAATCAACGATGCAAAAGTCAT 117

DB 4506 TTTATATATATGTGTTTGTATTTTAAAGTAAAAATATTTTGGAGAAATAGGGTTAGGGTTT 4565
QY 77 ATTGTTTTAAATCCAGATCGGAAATTTTGA AAAACATTTCTACTAAATTCGAAAT 136
DB 4566 AATTTTTTATAAGGTAGGGCGTGGAGAGACGATTTTATTTTTTATTTTCGTT 4625
QY 137 TAAATTTGAGCTTGCTTGATAAATGATAGTTGGCTCGGAGGTGATATTTTATCCTTGA 196
DB 4626 TTAATTTTGTATTCGATTTAGTTATGGGTTAAGTTTGGAGGGGACGTTGTTTTTAGA 4685
QY 197 TAGGAACACGCTCTTGAAGCTGTAAAGTTATCAATCGCTTGTGTTTCTATACAAAAAC 256
DB 4686 TAGGGTTTCGTTTATGTAAAGTGTATTTTGGTAAGGGATTTTTCGAGTAGAGGTAAC 4745
QY 257 ATGTTTT 263
DB 4746 GCGAATT 4752

RESULT 15
US-09-785-276A-8158
; Sequence 8158, Application US/09785276A
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007B
; CURRENT APPLICATION NUMBER: US/09/785,276A
; CURRENT FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8158
; LENGTH: 327
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 125, 132, 138, 143, 145, 150, 152, 156, 157, 165, 169, 180,
; LOCATION: 181, 182, 183, 184, 185, 186, 188, 190, 192, 193, 196, 204,
; LOCATION: 208, 212, 220, 223, 226, 228, 229, 232, 233, 234, 240, 244,
; LOCATION: 246, 250, 251, 254, 261, 272, 274, 276, 277, 289, 293
; OTHER INFORMATION: n = A,T,C or G
; NAME/KEY: misc.feature
; LOCATION: 296, 304
; OTHER INFORMATION: n = A,T,C or G
US-09-785-276A-8158

Query Match 6.5%; Score 45.2; DB 9; Length 327;
Best Local Similarity 41.7%; Pred. No. 5.9;
Matches 101; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

QY 16 TTTTTCCTTTTCTTACCTATTCAATAGTTCTTCTTGAATCAACGATGCAAAAGTCAT 75
DB 114 TTTTTCCTTTTCTTACCTATTCAATAGTTCTTCTTGAATCAACGATGCAAAAGTCAT 117
QY 76 TATTCCTTTTAAATCCAGATCGGAAATTTTGA AAAACATTTCTACTAAATTCGAAA 135
DB 114 TATTCCTTTTAAATCCAGATCGGAAATTTTGA AAAACATTTCTACTAAATTCGAAA 135

Db 118 TTTTNTTAAATNTAAAAAATANANATTTNTNTATNNCATTACTNAAANAGAAAAA 177
Qy 136 TTAAATTTGAGCTTGGCTTGATAAATGGTAGTGGCTCGGAGGTGATATTTTATCCCTTG 195
Db 178 AANNNNNNANANTNNATNTTTATATNTTANATTTATTAANGCNCNANNATNCNAAA 237
Qy 196 ATAGGAACACGCTTGAAGCTTAAGTTATTCAATCGCTTGTTTCTTCTATACAAAAA 255
Db 238 ANNATANAAAAANNNTANTAAACATNTATTATATTANANTNNGTGTNTTTTAAANAAAAANA 297
Qy 256 CA 257
Db 298 TA 299

Search completed: November 1, 2002, 23:31:39
Job time : 291.458 secs

ALL INFORMATION CONTAINED
HEREIN IS UNCLASSIFIED
DATE 07-11-2002 BY 60322 UCBAW

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 1, 2002, 23:16:04 ; Search time 1838.39 Seconds
(without alignments)
7331.704 Million cell updates/sec

Title: us-09-881-556-4
Perfect score: 623
Sequence: 1 cctcaatggtgaacatcaa.....aaaaaaaaaaaaaactcgag 623

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 21979536 seqs, 10817449327 residues

Total number of hits satisfying chosen parameters: 43959072

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending Patents_NA_Main.*
1: /cgn2_6/ptodata/1/pna/PCTUS_COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US06_COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US07_COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US080_COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US081_COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US082_COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US083_COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US084_COMB.seq.*
9: /cgn2_6/ptodata/1/pna/US085_COMB.seq.*
10: /cgn2_6/ptodata/1/pna/US086_COMB.seq.*
11: /cgn2_6/ptodata/1/pna/US087_COMB.seq.*
12: /cgn2_6/ptodata/1/pna/US088_COMB.seq.*
13: /cgn2_6/ptodata/1/pna/US089_COMB.seq.*
14: /cgn2_6/ptodata/1/pna/US090_COMB.seq.*
15: /cgn2_6/ptodata/1/pna/US091_COMB.seq.*
16: /cgn2_6/ptodata/1/pna/US092_COMB.seq.*
17: /cgn2_6/ptodata/1/pna/US093_COMB.seq.*
18: /cgn2_6/ptodata/1/pna/US094_COMB.seq.*
19: /cgn2_6/ptodata/1/pna/US095A_COMB.seq.*
20: /cgn2_6/ptodata/1/pna/US095B_COMB.seq.*
21: /cgn2_6/ptodata/1/pna/US095C_COMB.seq.*
22: /cgn2_6/ptodata/1/pna/US095D_COMB.seq.*
23: /cgn2_6/ptodata/1/pna/US096A_COMB.seq.*
24: /cgn2_6/ptodata/1/pna/US096B_COMB.seq.*
25: /cgn2_6/ptodata/1/pna/US096C_COMB.seq.*
26: /cgn2_6/ptodata/1/pna/US096D_COMB.seq.*
27: /cgn2_6/ptodata/1/pna/US096E_COMB.seq.*
28: /cgn2_6/ptodata/1/pna/US097A_COMB.seq.*
29: /cgn2_6/ptodata/1/pna/US097B_COMB.seq.*
30: /cgn2_6/ptodata/1/pna/US097C_COMB.seq.*
31: /cgn2_6/ptodata/1/pna/US098A_COMB.seq.*
32: /cgn2_6/ptodata/1/pna/US098B_COMB.seq.*
33: /cgn2_6/ptodata/1/pna/US098C_COMB.seq.*
34: /cgn2_6/ptodata/1/pna/US099A_COMB.seq.*
35: /cgn2_6/ptodata/1/pna/US099B_COMB.seq.*
36: /cgn2_6/ptodata/1/pna/US099C_COMB.seq.*
37: /cgn2_6/ptodata/1/pna/US100_COMB.seq.*
38: /cgn2_6/ptodata/1/pna/US101_COMB.seq.*
39: /cgn2_6/ptodata/1/pna/US6000_COMB.seq.*
40: /cgn2_6/ptodata/1/pna/US6001_COMB.seq.*
41: /cgn2_6/ptodata/1/pna/US6002_COMB.seq.*
42: /cgn2_6/ptodata/1/pna/US6003_COMB.seq.*
43: /cgn2_6/ptodata/1/pna/US6004_COMB.seq.*

44: /cgn2_6/ptodata/1/pna/US6005_COMB.seq.*
45: /cgn2_6/ptodata/1/pna/US6006_COMB.seq.*
46: /cgn2_6/ptodata/1/pna/US6007_COMB.seq.*
47: /cgn2_6/ptodata/1/pna/US6008_COMB.seq.*
48: /cgn2_6/ptodata/1/pna/US6009_COMB.seq.*
49: /cgn2_6/ptodata/1/pna/US6010_COMB.seq.*
50: /cgn2_6/ptodata/1/pna/US6011_COMB.seq.*
51: /cgn2_6/ptodata/1/pna/US6012_COMB.seq.*
52: /cgn2_6/ptodata/1/pna/US6013_COMB.seq.*
53: /cgn2_6/ptodata/1/pna/US6014_COMB.seq.*
54: /cgn2_6/ptodata/1/pna/US6015_COMB.seq.*
55: /cgn2_6/ptodata/1/pna/US6016_COMB.seq.*
56: /cgn2_6/ptodata/1/pna/US6017_COMB.seq.*
57: /cgn2_6/ptodata/1/pna/US6018_COMB.seq.*
58: /cgn2_6/ptodata/1/pna/US6019_COMB.seq.*
59: /cgn2_6/ptodata/1/pna/US6020_COMB.seq.*
60: /cgn2_6/ptodata/1/pna/US6021_COMB.seq.*
61: /cgn2_6/ptodata/1/pna/US6022_COMB.seq.*
62: /cgn2_6/ptodata/1/pna/US6023_COMB.seq.*
63: /cgn2_6/ptodata/1/pna/US6024_COMB.seq.*
64: /cgn2_6/ptodata/1/pna/US6025_COMB.seq.*
65: /cgn2_6/ptodata/1/pna/US6026_COMB.seq.*
66: /cgn2_6/ptodata/1/pna/US6027_COMB.seq.*
67: /cgn2_6/ptodata/1/pna/US6028_COMB.seq.*
68: /cgn2_6/ptodata/1/pna/US6029_COMB.seq.*
69: /cgn2_6/ptodata/1/pna/US6030_COMB.seq.*
70: /cgn2_6/ptodata/1/pna/US6031_COMB.seq.*
71: /cgn2_6/ptodata/1/pna/US6032_COMB.seq.*
72: /cgn2_6/ptodata/1/pna/US6033_COMB.seq.*
73: /cgn2_6/ptodata/1/pna/US6034_COMB.seq.*
74: /cgn2_6/ptodata/1/pna/US6035_COMB.seq.*
75: /cgn2_6/ptodata/1/pna/US6036_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	623	100.0	623	33	US-09-881-556-4 Sequence 4, Appli
2	57.8	9.3	787	29	US-09-739-449-2501 Sequence 2501, Ap
3	57.8	9.3	787	31	US-09-803-110-2501 Sequence 2501, Ap
4	54.2	8.7	543	29	US-09-739-449-8029 Sequence 8029, Ap
5	54.2	8.7	543	31	US-09-803-110-8029 Sequence 8029, Ap
6	54.2	8.7	1250	29	US-09-739-449-1859 Sequence 1859, Ap
7	54.2	8.7	1250	31	US-09-803-110-1859 Sequence 1859, Ap
8	53.2	8.5	562	33	US-09-873-402A-11512 Sequence 11512, A
9	53.2	8.5	562	59	US-60-209-830-11469 Sequence 11469, A
10	53	8.5	458	17	US-09-304-517A-251515 Sequence 251515,
11	53	8.5	458	17	US-09-371-146A-251515 Sequence 251515,
12	53	8.5	458	17	US-09-391-630-15771 Sequence 15771, A
13	53	8.5	458	25	US-09-654-617-63699 Sequence 63699, A
14	53	8.5	458	27	US-09-684-016-63699 Sequence 63699, A
15	53	8.5	458	36	US-60-144-084-5658 Sequence 251515,
16	53	8.5	458	53	US-09-985-678-251515 Sequence 251515,
17	52.6	8.5	466	22	US-09-565-306-29077 Sequence 29077, A
18	52.6	8.4	398	23	US-09-619-643-16189 Sequence 16189, A
19	52.6	8.4	411	25	US-09-654-617-340673 Sequence 340673,
20	52.6	8.4	411	27	US-09-684-016-340673 Sequence 340673,
21	52.6	8.4	1077	37	US-10-015-127-2817 Sequence 2817, Ap
22	52.6	8.4	1000	33	US-09-881-797-291 Sequence 291, App
23	52.2	8.4	1000	37	US-10-021-698-291 Sequence 291, App
24	52.2	8.4	413	24	US-09-637-086A-29308 Sequence 29308, A
25	52	8.3	413	24	US-09-637-086B-29308 Sequence 29308, A
26	52	8.3	413	24	US-09-739-449-2286 Sequence 2286, Ap
27	52	8.3	757	29	US-09-803-110-2286 Sequence 2286, Ap
28	52	8.3	757	31	US-09-803-110-2286 Sequence 2286, Ap
29	51.8	8.3	989	17	US-09-397-761A-2729 Sequence 2729, Ap
30	51.8	8.3	989	32	US-09-838-601-2729 Sequence 2729, Ap
31	51.6	8.3	664	29	US-09-739-449-739 Sequence 739, App

32	51.6	8.3	664	31	US-09-803-110-739	Sequence 739, App
c 33	51.4	8.3	637	25	US-09-654-617-429291	Sequence 429291, A
c 34	51.4	8.3	637	26	US-09-669-817A-12145	Sequence 12145, A
c 35	51.4	8.3	637	27	US-09-684-016-429291	Sequence 429291, A
c 36	51.4	8.3	634	29	US-09-739-449-2256	Sequence 2256, Ap
c 37	51.4	8.3	834	31	US-09-803-110-12256	Sequence 2256, Ap
c 38	51.2	8.3	1095	37	US-10-015-127-1938	Sequence 1938, Ap
c 39	51.2	8.2	530	33	US-09-873-402A-1650	Sequence 1650, Ap
c 40	51.2	8.2	530	59	US-60-209-830-1650	Sequence 1650, Ap
c 41	51.2	8.2	592	26	US-09-666-355A-7284	Sequence 7284, Ap
c 42	51.2	8.2	902	29	US-09-739-449-1929	Sequence 1929, Ap
c 43	51.2	8.2	902	31	US-09-803-110-1929	Sequence 1929, Ap
c 44	51	8.2	575	26	US-09-666-355A-12423	Sequence 12423, A
c 45	51	8.2	612	34	US-09-902-540-1357	Sequence 1357, Ap

ALIGNMENTS

```

RESULT 1
US-09-881-556-4
; Sequence 4, Application US/09881556
; GENERAL INFORMATION:
; APPLICANT: Simonsen, J. Neil
; TITLE OF INVENTION: Cryptosporidium Parvum Antigens, Antibodies Thereto and
; TITLE OF INVENTION: Diagnostic and Therapeutic Compositions Thereof
; FILE REFERENCE: 9000-0054
; CURRENT APPLICATION NUMBER: US/09/881.556
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/212,083
; PRIOR FILING DATE: 2000-06-15
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 623
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:C. Parvum AG2
US-09-881-556-4

```

Query Match	100.0%	Score 623;	DB 33;	Length 623;
Best Local Similarity	100.0%;	Pred. No. 3.3e-103;		
Matches 623;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	CCTCAAAATGGTCAACATCAAGAACTCTGTTCCAGACGAGCCCTATTACCTCACCAGAAA	60	
Db	1	CCTCAAAATGGTGAACATCAAGAACTCTGTTCCAGACGAGCCCTATTACCTCACCAGAAA	60	
QY	61	ACGAATCAAGTTCAAATCATCTTCTGTGACAGTTCCAGATCTGGATCAGTTCAAAATCT	120	
Db	61	ACGAATCAAGTTCAAATCATCTTCTGTGACAGTTCCAGATCTGGATCAGTTCAAAATCT	120	
QY	121	CCTTCGTGTACTATTCCAGAGACTGGATCAGACTCAGATCAGCGGCGTTGTGACAATTC	180	
Db	121	CCTTCGTGTACTATTCCAGAGACTGGATCAGACTCAGATCAGCGGCGTTGTGACAATTC	180	
QY	181	AGAGACTGGATCAGTTCAAATCATCTTCTGTCTACTATACCAGAAACAGGATCCAGCTCAG	240	
Db	181	AGAGACTGGATCAGTTCAAATCATCTTCTGTCTACTATACCAGAAACAGGATCCAGCTCAG	240	
QY	241	ATCAGACTCTGCTACTTCTCCAGAAGAGGATTGGACTCAGACGTTACCAATCACTTCT	300	
Db	241	ATCAGACTCTGCTACTTCTCCAGAAGAGGATTGGACTCAGACGTTACCAATCACTTCT	300	
QY	301	ACAGAACAACTCAAGCCAGCTACATATCCTTAACCAGAAAAATGAAATCATATAATTC	360	
Db	301	ACAGAACAACTCAAGCCAGCTACATATCCTTAACCAGAAAAATGAAATCATATAATTC	360	
QY	361	AGGAAGGTAATTCGAGTTTTTAATACACTAAATCTTCCAAATCAACCCAAATCTTTACGCA	420	
Db	361	AGGAAGGTAATTCGAGTTTTTAATACACTAAATCTTCCAAATCAACCCAAATCTTTACGCA	420	

Qy	421	AGCTGGCAGATGTGGAAGTTATGGGAAAGGATAAAATGGTTGATGGTGAGCAAGTAA	480
Db	421	AGCTGGCAGATGTGGAAGTTATGGGAAAGGATAAAATGGTTGATGGTGAGCAAGTAA	480
Qy	481	TCACATAAAATGACATATTCAAGATACTTCGAAAGAAATTAGAAACAAATGTAAAGTA	540
Db	481	TCACATAAAATGACATATTTCGAAGATACTTCGAAAGAAATTAGAAACAAATGTAAAGTA	540
Qy	541	TCTCATTGTATAAATATGGCCCTTAGCCATTTCCAAATATCTAAATTTGTCAACTCAAGTAA	600
Db	541	TCTCATTGTATAAATATGGCCCTTAGCCATTTCCAAATATCTAAATTTGTCAACTCAAGTAA	600
Qy	601	AAAAAAAAAAAAAAAAACTCGAG	623
Db	601	AAAAAAAAAAAAAAAAACTCGAG	623

RESULT 2

US-09-739-449-2501/c

; Sequence 2501, Application us/09739449

; GENERAL INFORMATION:

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; TITLE OR INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof

; FILE REFERENCE: 38-10(15490)C

; CURRENT APPLICATION NUMBER: US/09/739,449

; CURRENT FILING DATE: 2000-12-19

; PRIOR APPLICATION NUMBER: US 09/514,000

; PRIOR FILING DATE: 2000-02-23

; NUMBER OF SEQ ID NOS: 13351

; SEQ ID NO 2501

; LENGTH: 787

; TYPE: DNA

; ORGANISM: Agrobacterium tumefaciens

; FEATURE:

; NAME/KEY: unsure

; LOCATION: (1)..(787)

; OTHER INFORMATION: unsure at all n locations

US-09-739-449-2501

```

Query Match          9.3%; Score 57.8; DB 29; Length 787;
Best Local Similarity 50.2%; Pred. No. 0.82;
Matches 143; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY 333 AACCAAGAAAATGAAATCATATAATCAGGAAGGTAATTCGAGTGTTTTAATACACTAAAT 392
    || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 677 AAAAAATAAAATATAAAATATTTAAAAAANAATATAAAAAATATTAATTTAAAAAANAAT 618

QY 393 CTTCCTAAATCAACCCACTTTTCACGCAAGCTGGCAGATCTGGAAGAAGTTATGGGGAANAAG 452
    || |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 617 ATAAAAAATAATAAAAAAATATAAAAAATATAAAAAATATTTAAAAAANAATATAAAAAAATA 558

QY 453 GATAAAAATGGTTCGATGGTGAGCAAGTAATCACTAAAAAATGACATTATTGAAGATACTTCG 512
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 557 TATAAANAATATAAANAATAAANAATAAATAAANAATTAATAAAAAAANAATATTAATAAATAAAT 498

QY 513 AAAGAAATTAAGAACAAAAATGTAAGTATCTGCAATTGATAAATATGCGCCTTAGCCATTTTC 572
    |||| |||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 497 TAAAAAANAANAANAANAANAATATAAAAAATATAAAAAAANAANAATTAATAAAAAATNAAT 438

QY 573 CAATATCTAAATGTCACTCAAGTAAAAAANAANAANAANAANAANAANAANAANAANAANA 617
    |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 437 AAAAAATAAANAANAANAANAANAATNAANAANAANAANAANAANAANAANAANAANAANAANA 393

RESULT 3
US-09-803-110-2501/c
; Sequence 2501, Application US/09803110
; GENERAL INFORMATION:
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; TITLE OF INVENTION: Agrobacterium tumefaciens Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15490)D

```


Matches 152; Conservative 0; Mismatches 165; Indels 0; Gaps 0;

QY 301 ACAGAACAACTCAAGCCAGCTACATATCTTAACCAAGAAATGAAATCATATAATC 360
DB 403 ATATAATAATAAAAAATAGATAAAAAATATAAAAAAATAAAAAATAAAT 344
QY 361 AGGAAGGTAAATCGAGTTTAAATACACTAAATCTTCCAAATCAACCCATCTTTCACGCA 420
DB 343 AAAAAAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 284
QY 421 AGCTGCAGATGTCGAGTTTATGGGCAAGGATGATGATGTCGAGCAAGTAA 480
DB 283 AAAAAAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 224
QY 481 TCACATAAATGACATTTATGAGATCTTCCGAAGAAATAGAACCAAAATGTAAGTA 540
DB 223 TAAGATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 164
QY 541 TCTGCATTGATAATATGACCTTAGCCATTTCCAAATATCTAAATGTCACACTCAAGTAA 600
DB 163 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 104
QY 601 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 617
DB 103 TAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 87

RESULT 12
US-09-391-630-15771/c
; Sequence 15771, Application US/09391630
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Heck, Gregory R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Shukla, Hridayabhiramjam
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(15441)B
; CURRENT APPLICATION NUMBER: US/09/391,630
; CURRENT FILING DATE: 1999-09-08
; NUMBER OF SEQ ID NOS: 18889
; SEQ ID NO 15771
; LENGTH: 458
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: LIB3053-011-Q1-N1-C9
US-09-391-630-15771

Query Match 8.5%; Score 53; DB 17; Length 458;
Best Local Similarity 47.9%; Pred. No. 5.6;
Matches 152; Conservative 0; Mismatches 165; Indels 0; Gaps 0;

QY 301 ACAGAACAACTCAAGCCAGCTACATATCTTAACCAAGAAATGAAATCATATAATC 360
DB 403 ATATAATAATAAAAAATAGATAAAAAATATAAAAAAATAAAAAATAAAT 344
QY 361 AGGAAGGTAAATCGAGTTTAAATACACTAAATCTTCCAAATCAACCCATCTTTCACGCA 420
DB 343 AAAAAAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 284
QY 421 AGCTGCAGATGTCGAGTTTATGGGCAAGGATGATGATGTCGAGCAAGTAA 480
DB 283 AAAAAAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 224
QY 481 TCACATAAATGACATTTATGAGATCTTCCGAAGAAATAGAACCAAAATGTAAGTA 540
DB 223 TAAGATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 164
QY 541 TCTGCATTGATAATATGACCTTAGCCATTTCCAAATATCTAAATGTCACACTCAAGTAA 600

DB 163 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 104
QY 601 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 617
DB 103 TAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 87

RESULT 13
US-09-654-617-63699/c
; Sequence 63699, Application US/09654617
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 38-21(15097)D
; CURRENT APPLICATION NUMBER: US/09/654,617
; CURRENT FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 463173
; SEQ ID NO 63699
; LENGTH: 458
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: unsure at all n locations
US-09-654-617-63699

Query Match 8.5%; Score 53; DB 25; Length 458;
Best Local Similarity 47.9%; Pred. No. 5.6;
Matches 152; Conservative 0; Mismatches 165; Indels 0; Gaps 0;

QY 301 ACAGAACAACTCAAGCCAGCTACATATCTTAACCAAGAAATGAAATCATATAATC 360
DB 403 ATATAATAATAAAAAATAGATAAAAAATATAAAAAAATAAAAAATAAAT 344
QY 361 AGGAAGGTAAATCGAGTTTAAATACACTAAATCTTCCAAATCAACCCATCTTTCACGCA 420
DB 343 AAAAAAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 284
QY 421 AGCTGCAGATGTCGAGTTTATGGGCAAGGATGATGATGTCGAGCAAGTAA 480
DB 283 AAAAAAATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 224
QY 481 TCACATAAATGACATTTATGAGATCTTCCGAAGAAATAGAACCAAAATGTAAGTA 540
DB 223 TAAGATAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 164
QY 541 TCTGCATTGATAATATGACCTTAGCCATTTCCAAATATCTAAATGTCACACTCAAGTAA 600
DB 163 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 104
QY 601 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 617
DB 103 TAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 87

RESULT 14
US-09-684-016-63699/c
; Sequence 63699, Application US/09684016
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 38-21(15097)D
; CURRENT APPLICATION NUMBER: US/09/684,016
; CURRENT FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: US 09/654,617
; PRIOR FILING DATE: 2000-09-05
; NUMBER OF SEQ ID NOS: 463173
; SEQ ID NO 63699
; LENGTH: 458
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure

; LOCATION: (1)..(458)
; OTHER INFORMATION: unsure at all n locations
US-09-684-016-63699

Query Match 8.5%; Score 53; DB 27; Length 458;
Best Local Similarity 47.9%; Pred. No. 5.6;
Matches 152; Conservative 0; Mismatches 165; Indels 0; Gaps 0;

QY 301 ACAGAACAACTCAAGCCAGCTACATATCTTCCAAATCAACCAAGAAATGAAAAATCATATATATC 360
DB 403 ATATAATAATAAAAAATAGATAAAAAATATAAAAAAATAAAAAAATAAAAAATAAAT 344
QY 361 AGGAAGGTAATTCGAGTTTAAATACACTAAATCTTCCAAATCAACCAATCTTTCACGCA 420
DB 343 AAAAAAATAATAAAATTAATAAAAAAATAAAAAAATAAAAAAATAAAAAATAA 284
QY 421 AGCTGGCAGATGTGGAAGTTTGGGGAAGAGTAATAAATGGTTGATGGTGAGCAAGTAA 480
DB 283 AAAAAAATAATAAAATTAATAAAAAATATAAATATAAATAAATAAATAAATAAATA 224
QY 481 TCACATAAAATGACATTATTGAGATATCTCGGAAGAAATAGAAACAAATGTAAAGTA 540
DB 223 TAGATAATAAATAAATAAAAAAATAAATAATAAATAAATAAATAAATAAATAAATA 164
QY 541 TCTGATTGATAATATGGCCTTAGCCATTCCCAATATCTAAATTTGTCACACTCAAGTAA 600
DB 163 AAATTAAAAAATAAAAAAATAAATAATAAATAAATAAATAAATAAATAAATAAATA 104
QY 601 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 617
DB 103 TAAAAAATAATATATAA 87

RESULT 15

US-09-985-678-251515/c
; Sequence 251515, Application US/09985678
; GENERAL INFORMATION:
; APPLICANT: Cheikh, Nordine
; TITLE OF INVENTION: Annotated plant Genes
; FILE REFERENCE: 16517.255/38-21(15097)F
; CURRENT APPLICATION NUMBER: US/09/985,678
; CURRENT FILING DATE: 2001-11-05
; PRIOR APPLICATION NUMBER: US 09/304,517
; PRIOR FILING DATE: 1999-05-06
; NUMBER OF SEQ ID NOS: 295529
; SEQ ID NO 251515
; LENGTH: 458
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(458)
; OTHER INFORMATION: unsure at all n locations
US-09-985-678-251515

Query Match 8.5%; Score 53; DB 36; Length 458;
Best Local Similarity 47.9%; Pred. No. 5.6;
Matches 152; Conservative 0; Mismatches 165; Indels 0; Gaps 0;

QY 301 ACAGAACAACTCAAGCCAGCTACATATCTTACCAAGAAATGAAAAATCATATATATC 360
DB 403 ATATAATAATAAAAAATAGATAAAAAATATAAAAAAATAAAAAAATAAAAAATAAAT 344
QY 361 AGGAAGGTAATTCGAGTTTAAATACACTAAATCTTCCAAATCAACCAATCTTTCACGCA 420
DB 343 AAAAAAATAATAAAATTAATAAAAAAATAAAAAAATAAAAAAATAAAAAATAA 284
QY 421 AGCTGGCAGATGTGGAAGTTTATCGGGAAGAGTAATAAATGGTTGATGGTGAGCAAGTAA 480
DB 283 AAAAAAATAAATAAATAAAAAAATAAATAATAAATAAATAAATAAATAAATAAATA 224
QY 481 TCACATAAAATGACATTATTGAGATATCTCGGAAGAAATAGAAACAAATGTAAAGTA 540

DB 223 TAAGATAATAAATAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAATA 164
QY 541 TCTGCATTGATAAATAATATGGCCTTAGCCATTTCACAAATATCTAAATTTGTCACACTCAAGTAA 600
DB 163 AAATTAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAATAA 104
QY 601 AAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 617
DB 103 TAAAAAATAATATATAA 87

Search completed: November 2, 2002, 00:37:09
Job time : 1846.39 secs

THIS PAGE BLANK (USPTO)

QY 591 ACTCAAGTAAAAA 616
Db 500 NAAANNNAAAAA 525

RESULT 6
US-10-155-881-3856
; Sequence 3856, Application US/10155881
; GENERAL INFORMATION:
; APPLICANT: Dolson, Stanton B.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: Lutfiyya, Linda L.
; APPLICANT: McIninch, James
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE OF INVENTION: TRANSCRIPTION IN PLANTS
; FILE REFERENCE: 38-21(15300)J
; CURRENT APPLICATION NUMBER: US/10/155,881
; CURRENT FILING DATE: 2002-05-22
; NUMBER OF SEQ ID NOS: 37595
; SEQ ID NO 3856
; LENGTH: 2077
; TYPE: DNA
; ORGANISM: Zee mayes
US-10-155-881-3856

Query Match 7.5%; Score 47; DB 12; Length 2077;
Best Local Similarity 46.7%; Pred. No. 3.9;
Matches 149; Conservative 0; Mismatches 170; Indels 0; Gaps 0;
QY 299 CTACAGACAACTCAAGCCAGCTACATATCTCAACCAAGAAATGAAATCATATAA 358
Db 1384 CTAATATAGAGACATAAGACACCCCTTAATGTAATGAAGTAACGACAAATGAAGCTCA 1443
QY 359 TCAGGAAGGTAATTCGAGTTTAAATACACTAAATCTTCCAAATCAACCCCAATCTTTCACG 418
Db 1444 GATGATATATCTACATATCTTCGTAAAAA 1503
QY 419 CAAGCTGGCAGATGTGGAAAGTTATGGGAAAGGATATAATGTTGATGGTGAGCAAGT 478
Db 1504 AAGGGGAAAAA 1563
QY 479 AATCACTAAAAATGACATTATCAAGTACTTCGAAGAAATTTAGAACAAATGTAAAG 538
Db 1564 AAAAAA 1623
QY 539 TATCTGATTGATAATATATGGCCTTAGCCATTTCCAAATATCTAAATTTGTCAACTCAAGT 598
Db 1624 AAAAAA 1683
QY 599 AAAAAA 617
Db 1684 AAAAAA 1702

RESULT 7
US-10-172-086-112/c
; Sequence 112, Application US/10172086
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Method and nucleic acids for the differentiation
; FILE OF INVENTION: of prostate tumors
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/172,086
; CURRENT FILING DATE: 2002-06-13
; NUMBER OF SEQ ID NOS: 116
; SEQ ID NO 112
; LENGTH: 6944
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-172-086-112

Query Match 7.5%; Score 46.8; DB 12; Length 6944;
Best Local Similarity 47.9%; Pred. No. 4.8;
Matches 135; Conservative 0; Mismatches 147; Indels 0; Gaps 0;
QY 333 AACCAAGAAATGAAATCATATAATCAAGAGGTAAATTCGAGTTTAAATACACTAAAT 392
Db 5782 AAAAAA 5723
QY 393 CTTCCAAATCAACCCCAATCTTTCACGCAAGCTGCGAGATCTGGAAAGTTATGGGAAAG 452
Db 5722 TTCCAAAAA 5663
QY 453 GATAAAATGTTGATGGTGAGCAAGTAAATCACTAAAAATGACATTTATTTGAAGATACCTCG 512
Db 5662 AAAAAA 5603
QY 513 AAGAAATTTAGAACCAAAATGTAAGTATCTGCAATTCATATAATATGCCCTTACCCATTTTC 572
Db 5602 AAAAAA 5543
QY 573 CAAATATCTAAATTTGCAACTCAAGTAAAAA 614
Db 5542 AAAAAA 5501

RESULT 8
US-10-155-881-13051
; Sequence 13051, Application US/10155881
; GENERAL INFORMATION:
; APPLICANT: Dolson, Stanton B.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jingdong
; APPLICANT: Lutfiyya, Linda L.
; APPLICANT: McIninch, James
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE OF INVENTION: TRANSCRIPTION IN PLANTS
; FILE REFERENCE: 38-21(15300)J
; CURRENT APPLICATION NUMBER: US/10/155,881
; CURRENT FILING DATE: 2002-05-22
; NUMBER OF SEQ ID NOS: 37595
; SEQ ID NO 13051
; LENGTH: 2606
; TYPE: DNA
; ORGANISM: Glycine max
US-10-155-881-13051

Query Match 7.5%; Score 46.6; DB 12; Length 2606;
Best Local Similarity 46.7%; Pred. No. 4.8;
Matches 148; Conservative 0; Mismatches 169; Indels 0; Gaps 0;
QY 301 ACAGACAAACTCAAGCCAGCTACATATCTTAACCAAGAAATGAAATCATATAATTC 360
Db 1767 AAAAAA 1826
QY 361 AGGAAGTAAATTCGAGTTTAAATACACTAAATCTTCCAAATCAACCCCAATCTTTTCACGCA 420
Db 1827 AAAAAA 1886
QY 421 AGCTGGCAGATGTGGAAAGTTATGGGAAAGGATATAATTCGTTGATGGTGAGCAAGTAA 480
Db 1887 AAAAAA 1946
QY 481 TCACTAAAAATGACATTTATTTGAAGTACTTTCGAAAGAAATTTAGAACAAATTTGTAAGTA 540
Db 1947 AAAAAA 2006
QY 541 TCTGCATTGATAATATGCGCTTAGCCATTTCCAAATATCTAAATTTGCTCAACTCAAGTAA 600
Db 2007 AAAAAA 2066
QY 601 AAAAAA 617

Db 2067 AAAAAAAAAAAAAAAAAA 2083

RESULT 9

US-10-221-613-93/c

; Sequence 93, Application US/10221613

; GENERAL INFORMATION:

; APPLICANT: OLEK, Alexander

; APPLICANT: PIEPENBROCK, Christian

; APPLICANT: BERLIN, Kurt

; TITLE OF INVENTION: Diagnosis of Diseases Associated with Cell Cycle

; CURRENT APPLICATION NUMBER: US/10/221,613

; PRIOR FILING DATE: 2002-09-13

; PRIOR APPLICATION NUMBER: PCT/EP01/02945

; DE 10013847.00

; DE 10019058.8

; DE 10019173.8

; DE 10032529.7

; DE 10043826.1

; PRIOR FILING DATE: 2001-03-15

; 2000-03-15

; 2000-04-06

; 2000-04-07

; 2000-06-30

; 2000-09-01

; NUMBER OF SEQ ID NOS: 428

; SEQ ID NO 93

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-221-613-93

Query Match 7.4%; Score 46.4; DB 13; Length 4661;

Best Local Similarity 44.9%; Pred. No. 5.6; Indels 0; Gaps 0;

Matches 176; Conservative 0; Mismatches 216;

QY 224 AACAGGATCCAGCTCAGATCAGATCTCTGCTACTCTCCAGAGAGGATTGGACTCAGAA 283

Db 4134 AAAAAAAAAACAAATAATACATAATCCAACTCAATCACTCCGATCCAAAC 4075

QY 284 CTTTACCATCACTTCTACAGAACAACTCAAGCCAGCTACATCTCAACCAAGAAAA 343

Db 4074 CCATCTCAATCCCTCCAAAAAATACTTTAATCTACCCCTTATTACAAATAACAAT 4015

QY 344 TGAAATCATATAATCAGGAGGTAATTCGAGTTTAAATACACTAAATCTTCCAAATCA 403

Db 4014 CAACCTCTCAAAATCAAAATCCCTTAACTAAATAATTAACCTAAACACACCTTAATAA 3955

QY 404 ACCCAATCTTTCACGCAAGCTGGCAGATGTGGAAAGTTATGGGAAAGGATAAAATGGT 463

Db 3954 AAACCACTAAATAAATAAACTTATAACCTCAAAACACTATAAACACAAAAAACCTAA 3895

QY 464 TGATGGTGAGCAAGTAATCACTAAATAATGACATTAATGGAATACCTTCGAAAGAAATAG 523

Db 3894 CTCCTCCGAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 3835

QY 524 AAACAAATGTAAAGTATCTGATTAATAATATGGCTTAGCCATTTCCTCAATATCTAA 583

Db 3834 AAAAAAAAAAAAAAACAATAATTTCTAAAAAATAAATAAATAAATAAATAAATAAATAA 3775

QY 584 ATTGCACTCAAGTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 615

Db 3774 AATAAACCAACAAACAAACGACCAACCAAAA 3743

RESULT 10

US-10-239-676-165/c

; Sequence 165, Application US/10239676

; GENERAL INFORMATION:

; APPLICANT: OLEK, Alexander

; APPLICANT: PIEPENBROCK, Christian

; APPLICANT: BERLIN, Kurt

; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation

; FILE REFERENCE: 5013.1003

; CURRENT APPLICATION NUMBER: US/10/239,676

; CURRENT FILING DATE: 2002-09-24

; PRIOR APPLICATION NUMBER: PCT/EP01/03968

; DE 10019058.8

; DE 10019173.8

; DE 10032529.7

; DE 10043826.1

; PRIOR FILING DATE: 2001-04-06

; 2000-04-06

; 2000-06-30

; 2000-09-01

; NUMBER OF SEQ ID NOS: 228

; SEQ ID NO 165

; LENGTH: 13606

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

; NAME/KEY: unsure

; LOCATION: (4649, 7634, 7657)

US-10-239-676-165

Query Match 7.4%; Score 46.4; DB 13; Length 13606;

Best Local Similarity 47.0%; Pred. No. 6.1; Indels 0; Gaps 0;

Matches 143; Conservative 0; Mismatches 161;

QY 314 AAAGCAGCTCATATCTCTCAACCAAGAAAATGAAAATCATATATATCAGGAAGGTAATTC 373

Db 11279 AAAATTACTACAAACCTTAAAAACATACCTTAAAAAATAAATAAATAAATAAATAA 11220

QY 374 GAGTTTTTAACACTAAATCTTCCAAATCAACCCAAATCTTTACGCAAGCTGGCAGATGT 433

Db 11219 AAAAAAAAAAAAAATATTCTCCAAACACCCCAAAACCTAAACCTTAACCCAAAC 11160

QY 434 GGAAGTTATGGGAAAAGGATAAATGTTGATGGTGAGCAAGTAATCACTAAAAATGA 493

Db 11159 AAACTACAAAAACAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 11100

QY 494 CATATTGGAAGTACTTCGAAAGAAATTTAGAAACAAAAATGTAAAGTATCTGCATTGATA 553

Db 11099 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 11040

QY 554 ATATGGCCCTTAGCCATTTCCAAATATCTAAATTTGTCAACTCAAGTAAAAAATAAATA 613

Db 11039 TTATATATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 10980

QY 614 AAAA 617

Db 10979 AAAA 10976

RESULT 11

US-10-155-881-17885

; Sequence 17885, Application US/10155881

; GENERAL INFORMATION:

; APPLICANT: Dotson, Stanton B.

; APPLICANT: Kovalic, David K.

; APPLICANT: Liu, Jingdong

; APPLICANT: Lutfiyya, Linda L.

; APPLICANT: McIninch, James

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH

; FILE REFERENCE: 38-21(15300)J

; CURRENT APPLICATION NUMBER: US/10/155,881

; CURRENT FILING DATE: 2002-05-22

; NUMBER OF SEQ ID NOS: 37595

; SEQ ID NO 17885


```
; LENGTH: 2729
; TYPE: DNA
; ORGANISM: Glycine max
US-10-155-881-17885

Query Match      7.4%  Score 46.2;  DB 12;  Length 2729;
Best Local Similarity 46.0%;  Pred. No. 5.8;  Mismatches 0;  Gaps 0;
Matches 156;  Conservative 0;

Qy 279 CAGAACGTTACCAATCACTCTCAGACAACTCAAGCCAGCTACATATCCCTAACCAA 338
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2125 CAGCAAGTTATCAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 2184
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 339 GAAATGAAATCATATAATCAGGAGGTAATTCGAGCTTTTAATACACTAAATCTTCCA 398
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2185 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 2244
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 399 AATCAACCCCAATCTTTCACGCAAGCTGCAGATGTGGAAGCTTATGGGAAAGATAAA 458
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2245 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 2304
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 459 ATGCTTGATGTGAGCAAGTAATCACTAAATGACATTTATGAAGATACCTCGAAAGAA 518
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2305 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 2364
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 519 ATTGAACAAATGTAAGTATCTGCATTTGATTAATATGCGCTTAGCCATTTCCAAATA 578
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2365 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 2424
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 579 TCTAAATTTGCTCACTCAAGTAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 617
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2425 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 2463
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 12
US-10-198-846-1483
; Sequence 1483, Application US/10198846
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; PRIOR FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1483
; LENGTH: 525
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 46, 52, 75, 78, 80, 85, 91, 100, 102, 107, 109, 114,
; LOCATION: 123, 127, 136, 142, 149, 156, 159, 160, 162, 168, 169, 173,
; LOCATION: 177, 195, 203, 207, 210, 211, 213, 219, 220, 224, 230, 251,
; LOCATION: 261, 268, 277, 309, 327, 333, 337, 339, 355, 359, 364
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 365, 380, 389, 399, 404, 408, 413, 416, 429, 434, 435, 438,
; LOCATION: 440, 458, 473, 479, 490, 511, 519, 520, 522, 525
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-1483

Query Match      7.3%  Score 45.6;  DB 13;  Length 525;
Best Local Similarity 39.1%;  Pred. No. 6.5;  Mismatches 0;  Gaps 0;
Matches 156;  Conservative 0;

Qy 310 ACTCAAGCCAGCTACATATCTTACCAAGAAATGAAATCAATTAATTAATCAGGAAGTA 369
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2293 ATTAACATCTACTATATATTCTTAACACTATTCCAAAACAAATACAAATTAACAAACA 2234
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 370 ATTCGAGTTTTTAATACATAAATCTTCCAAATCAACCAATCTTTCAGGCAAGCTGCAG 429
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2233 AACCAAAAAAACAATAAATAATACACCTCGAATACACACACACACACACACACAC 2174
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Qy 219 CCAGAAACAGGATCCAGCTCAGATCAGACTCTGCTACTTCTCCAGAAAGAGATTGGACT 278
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 38 CCAACCGCNGTGGCGGCCGAAAGTAACATTTTTTTTTTTNNNNNNNNNNNNNNNNNN 97
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 279 CAGAACGTTACCAATCACTCTCAGACAACTCAAGCCAGCTACATATCCCTAACCAA 338
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 98 CANANAGAAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAAN 157
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 339 GAAATGAAATCATATAATCAGGAGGTAATTCGAGCTTTTAATACACTAAATCTTCCA 398
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 158 ANNANAAAAAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA 217
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 399 AATCAACCCCAATCTTTCACGCAAGCTGCAGATGTGGAAGCTTATGGGAAAGATAAA 458
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 218 ANNANAAAAAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA 277
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 459 ATGCTTGATGTGAGCAAGTAATCACTAAATGACATTTATGAAGATACCTTCGAAAGAA 518
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 278 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 337
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 519 ATTGAACAAATGTAAGTATCTGCATTTGATTAATATGCGCTTAGCCATTTCCAAATA 578
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 338 ANAACCAAAAAAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNAANNA 397
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 579 TCTAAATTTGCTCACTCAAGTAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 617
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 398 TTTAAANNAANNAATNAANTAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 436
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
US-10-221-613-311/c
; Sequence 311, Application US/10221613
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Cell Cycle
; FILE REFERENCE: 5013.1004
; CURRENT APPLICATION NUMBER: US/10/221.613
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: PCT/EP01/02945
; DE 10013847.00
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-03-15
; 2000-03-15
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 428
; SEQ ID NO 311
; LENGTH: 7312
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-221-613-311

Query Match      7.3%  Score 45.6;  DB 13;  Length 7312;
Best Local Similarity 46.8%;  Pred. No. 8.3;
Matches 144;  Conservative 0;  Mismatches 164;  Indels 0;  Gaps 0;

Qy 310 ACTCAAGCCAGCTACATATCTTACCAAGAAATGAAATCAATTAATTAATCAGGAAGTA 369
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2293 ATTAACATCTACTATATATTCTTAACACTATTCCAAAACAAATACAAATTAACAAACA 2234
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 370 ATTCGAGTTTTTAATACATAAATCTTCCAAATCAACCAATCTTTCAGGCAAGCTGCAG 429
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2233 AACCAAAAAAACAATAAATAATACACCTCGAATACACACACACACACACACACAC 2174
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Qy	430	ATGTGGAAAGCTTATGGGGAAGAGATAAAATGGTTGATGGTGAGCAAGTAATCACTAAAA	489
Db	2173	ACAAAAAACACCATTTAAAAATAAATTAATCTACTTTACAATACACCAAAATTTAACAAAAAT	2114
Qy	490	ATGACATATTGAAGATACTTCGAAGAATAATTAGAAACAATAATGTAAAGTATCTGCATTG	549
Db	2113	AATAAAATTAATACTAAACAACATAATTTAATAATAAAATATAAAAAAACTTACTTCCAAA	2054
Qy	550	ATAAATATGGCCTTAGCCATTTCCAAATATCTAAATTTGTCAACTCAAGTAAAAAATAAAAA	609
Db	2053	AAATATTTAACTAAANAACCTCATAAAAAATAAAAAAATCCCGNACAATACACCTAAA	1994
Qy	610	AAAAAAA 617	
Db	1993	TAAAAAAA 1986	
RESULT 14			
US-10-239-676-211/c			
; Sequence 211, Application US/10239676			
; GENERAL INFORMATION:			
; APPLICANT: OLEK, Alexander			
; APPLICANT: PIEPENBROCK, Christian			
; APPLICANT: BERLIN, Kurt			
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation			
; FILE REFERENCE: 5013.1003			
; CURRENT APPLICATION NUMBER: US/10/239,676			
; CURRENT FILING DATE: 2002-09-24			
; PRIOR APPLICATION NUMBER: PCT/EP01/03968			
; DE 10019058.8			
; DE 10019173.8			
; DE 10032529.7			
; DE 10043826.1			
; PRIOR FILING DATE: 2001-04-06			
; 2000-04-06			
; 2000-04-07			
; 2000-06-30			
; 2000-09-01			
; NUMBER OF SEQ ID NOS: 228			
; SEQ ID NO 211			
; LENGTH: 8996			
; TYPE: DNA			
; ORGANISM: Artificial Sequence			
; FEATURE:			
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)			
US-10-239-676-211			

Db 6576 ATAAAAATATAAAAAAAAAAAAAAAAAATA 6546

RESULT 15

```

US-10-239-676-117/c
; Sequence 117, Application US/10239676
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
; FILE REFERENCE: 5013.1003
; CURRENT APPLICATION NUMBER: US/10/239,676
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: PCT/EP01/03968
; DE 10019058.8
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIOR FILING DATE: 2001-04-06
; 2000-04-06
; 2000-04-07
; 2000-06-30
; 2000-06-30
; 2000-09-01
; NUMBER OF SEQ ID NOS: 228
; SEQ ID NO 117
; LENGTH: 11036
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; NAME/KEY: unsure
; LOCATION: (1927..1928, 1943, 1956, 1958, 1981)
US-10-239-676-117

Query Match          7.2%; Score 45; DB 13; Length 11036;
Best Local Similarity 47.4%; Pred. No. 11;
Matches 135; Conservative 0; Mismatches 150; Indels 0; Gaps

QY 333 AACCAAGAAATGAAATCATATAATAACGAGGTAATTCGAGTTTATACACTAAAT 392
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 7784 AACTATAAAATTCNAACTCTATAACCATATAATAATAATAAATACTAAATATCAAAAC 7725
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 393 CTTCCAAATCAOCCAACTTTTCACGCAAGTGGCAGATGTGGAAGTTATGGGGAAG 452
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 7724 TATAAAAAATAACCAATTTATTATACATATTTTACACAAAAATTAATAATTTTTTAAAAA 7665
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 453 GATAAATGTTGATGGTGAGCAAGTAATCACTAAAAATGCATTTATGAAGATACTCG 512
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 7664 AAAATTAAATATTACCACCTTTAAAAAAATCTATAAAAAATACAAAAAACACATTACAATA 7605
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 513 AAAGAATATTAGAACAAAAATGTAAGATATCTGCAATTGATAAATATGCGCTTAGCCATTTC 572
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 7604 ACACTAACTACTAATCAACACTACATATAATATATTTTAAATAATAAATAATTCTAA 7545
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 573 CAATATCTAAATGTGCACTCAAGTAAAAAATAAAAAA 617
    ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
Db 7544 AAATATTTAAATTTTAAATTTTATAAAAATAAATAAATAAAAAA 7500
    ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||

```

Search completed: November 1, 2002, 23:32:35
Job time : 298.542 secs